

# Empirische Linguistik im Recht – am Beispiel des Wandels des Staatsverständnisses im Sicherheitsrecht, öffentlichen Wirtschaftsrecht und Sozialrecht der Schweiz

Empirical linguistics in law, taking as an example changes in our understanding of the “state” in security law, public commercial law and social law in Switzerland

Andreas Abegg/Noah Bubenhofer

Translated by Alison Lewis

## I. Wandel des Staates als Recht-Fertigung

Regelmässig wird zunächst in deskriptiver Weise dargelegt, dass sich der Staat (oder alternativ: die Verwaltung oder das Verhältnis von Staat und Gesellschaft) gewandelt habe. Gestützt auf diesen Befund wird sodann implizit oder explizit gefolgert, dass der neuen Wirklichkeit eine Anpassung des Rechts zu folgen habe. Während die einen etwa den Wandel des Staates als Abkehr vom autoritären Staat hin zu einem gewährleistenden bürgernahen Staat beschreiben,<sup>1</sup> findet sich andernorts ein Staatsverständnis, das (wieder) die Verantwortung des Staates gegenüber seiner Gesellschaft in den Vordergrund stellt.<sup>2</sup>

## I. Transformation of the state as justification

Generally speaking, we start with a descriptive presentation of the fact that the state (or alternatively the administration, or the relationship between the state and society) has changed. On the basis of this finding, we conclude (implicitly or explicitly) that the law must be amended to take account of the new reality. While some will perhaps describe the change in terms of the state turning away from the authoritarian model and moving towards a citizen-centred, guarantee-providing model,<sup>1</sup> elsewhere we find an understanding of the state that (once again) places the main emphasis on the state's responsibility towards society.<sup>2</sup>

---

1 Siehe unter anderen: Hans Rudolf Trüeb, Grundversorgung mit Fernmeldediensten: Gemeindeförderung oder Mittel der Strukturerhaltung?, Aktuelle Juristische Praxis (AJP) (2002), 1186–1194; Andreas Lienhard, PPP im Verfassungsrecht, in: Andreas Lienhard/Thomas Pfisterer (Hrsg.), PPP – Was fehlt zum Durchbruch (2010), 17–36; Urs Bolz, Public Private Partnership (PPP) in der Schweiz, Zentralblatt für Staats- und Verwaltungsrecht (ZBl) (2004), 561–596.

2 Vgl. z.B. Markus Müller, Verwaltungsrecht: Eigenheit und Herkunft (2006); Markus Müller/Christoph Jenni, Die polizeiliche Generalklausel – Ein Institut mit Reformbedarf, Polizei & Sicherheit (2008), 4–18; Markus Müller/Christoph Jenni, Notrecht ... abermals zur polizeilichen Generalklausel, Polizei & Sicherheit (2010), 101–109.

---

1 See also: Hans Rudolf Trüeb, Grundversorgung mit Fernmeldediensten: Gemeindeförderung oder Mittel der Strukturerhaltung?, Aktuelle Juristische Praxis (AJP) (2002), 1186–1194; Andreas Lienhard, PPP im Verfassungsrecht, in: Andreas Lienhard/Thomas Pfisterer (Hrsg.), PPP – Was fehlt zum Durchbruch (2010), 17–36; Urs Bolz, Public Private Partnership (PPP) in der Schweiz, Zentralblatt für Staats- und Verwaltungsrecht (ZBl) (2004), 561–596.

2 Cf. eg. Markus Müller, Verwaltungsrecht: Eigenheit und Herkunft (2006); Markus Müller/Christoph Jenni, Die polizeiliche Generalklausel – Ein Institut mit Reformbedarf, Polizei & Sicherheit (2008), 4–18; Markus Müller/Christoph Jenni, Notrecht ... abermals zur polizeilichen Generalklausel, Polizei & Sicherheit (2010), 101–109.

Bei derartigen Verweisen (sei es von Seiten der Rechtsdogmatik oder der Rechtsgeschichte<sup>3</sup>) auf den Wandel des Staates geht es zumeist um eine „Recht-Fertigung“ (Wiethölter)<sup>4</sup> normativer Aussagen; die jeweils vorgeschlagenen normativen Veränderungen werden mit einem veränderten Verständnis von Staatlichkeit „be-gründet“. Unser Recht liegt in der Sprache<sup>5</sup>: Normative Folgen werden mit Sprache umschrieben, und diese imperative Anweisung durch das Recht erhält erst durch die Be-gründung, dass bestimmte Handlungsanweisungen den gemeinsamen Vorstellungen und Zielen dienen, ihre Legitimation und Rechtmässigkeit – ihren Grund und Boden.<sup>6</sup> Damit hängt aber die Qualität der Recht-Fertigung vom Grund ab, auf welchem diese steht. In Bezug auf die vorliegende Fragestellung: Hat sich die Staatlichkeit *tatsächlich* verändert – und wenn ja, auf welche Weise?

Such references (emanating either from legal doctrine or from the history of law<sup>3</sup>) to our changing view of the state are mostly a “justification (Recht-Fertigung = law production)” (Wiethölter)<sup>4</sup> of normative statements; the proposed normative changes are “justified (be-gründet = established)” by a change in our understanding of statehood. Our law is based in language<sup>5</sup>: normative consequences are circumscribed with language, and the imperative dispositions provided by the law only attain their true basis – their legitimacy and lawful nature – through the justification (Be-gründung) that certain dispositions concerning our actions serve common preconceptions and aims.<sup>6</sup> This means, however, that the quality of the justification (Recht-Fertigung) depends on the basis on which the justification stands, i.e. in regard to the question we are considering here: has statehood *really* changed – and if so, in what ways?

3 Dies gilt zumindest für eine zunehmend als praxisrelevant verstandene Rechtsgeschichte: Jörn Eckert, Die Krise der Rechtsgeschichte und die Frage nach ihrem Nutzen für die Theorie und die Praxis des Rechts, in: Jörn Eckert/Hans Hattenhauer (Hrsg.), Der praktische Nutzen der Rechtsgeschichte. Hans Hattenhauer zum 8. September 2001 (2003), 121–158, 130f.

4 Rudolf Wiethölter, Recht-Fertigungen eines Gesellschafts-Rechts, in: Christian Joerges/Gunther Teubner (Hrsg.), Rechtsverfassungsrecht: Recht-Fertigung zwischen Privatrechtsdogmatik und Gesellschaftstheorie (2003), 13–21.

5 Diese Formulierung knüpft an Heidegger an: Martin Heidegger, Der Weg zur Sprache, in: Unterwegs zur Sprache (1985), 228–257.

6 Zur Legitimation von Recht durch Begründung sei stellvertretend für viele auf Robert Cover's Analysen des vorstaatlichen Rechts verwiesen: Robert M. Cover, Nomos and Narrative, Harvard Law Review (1983), 4–67.

3 This applies at least in regard to a history of the law which is increasingly understood as relevant to practice: Jörn Eckert, Die Krise der Rechtsgeschichte und die Frage nach ihrem Nutzen für die Theorie und die Praxis des Rechts, in: Jörn Eckert/Hans Hattenhauer (Hrsg.), Der praktische Nutzen der Rechtsgeschichte. Hans Hattenhauer zum 8. September 2001 (2003), 121–158, 130f.

4 Rudolf Wiethölter, Recht-Fertigungen eines Gesellschafts-Rechts, in: Christian Joerges/Gunther Teubner (Hrsg.), Rechtsverfassungsrecht: Recht-Fertigung zwischen Privatrechtsdogmatik und Gesellschaftstheorie (2003), 13–21.

5 This formulation is linked to Heidegger in: Martin Heidegger, Der Weg zur Sprache, in: Unterwegs zur Sprache (1985), 228–257.

6 Concerning the legitimacy of law through justification, reference is made (representing many authors) to Robert Cover's analyses of the law antecedent to the state: Robert M. Cover, Nomos and Narrative, Harvard Law Review (1983), 4–67.

## II. Empirie als erweiterte hermeneutische Be-Gründung

Die Jurisprudenz befasst sich mit dem Verstehen des normativen Gehalts sprachlicher Äusserungen. Dieses Verstehen ist nach Gadamer stets<sup>7</sup> Auslegung, mit welcher der Auslegende den Sinn des Textes zu verstehen sucht. Mit seinem eigenen Erfahrungshintergrund analysiert der Auslegende jene bedeutsamen Umstände, welche zur sprachlichen Äusserung geführt haben. Es geht somit nicht nur darum, dass eine normative Aussage richtig auf den zur Beurteilung stehenden Lebenssachverhalt angewandt wird. Vielmehr sind nach der *Hermeneutik* beide – Lebenssachverhalt und sprachliche Äusserung – aufeinander angewiesen.<sup>8</sup> Gadamer spricht hierbei von einer „ontologischen Wendung“<sup>9</sup>; das in der Tradition von Descartes, Kant und Fichte stehende Subjekt ersetzt sich durch eine Sprachtheorie. Diese bleibt aber der subjektphilosophischen Tradition insofern treu, als das Verstehen dem Vorverständnis der Verstehenden ausgesetzt bleibt.<sup>10</sup>

## II. Empiricism as expanded hermeneutic justification (Be-Gründung)

Jurisprudence is concerned with understanding the normative content of linguistic statements. According to Gadamer the act of understanding is always<sup>7</sup> an act of interpretation, whereby the person doing the interpreting seeks to understand the meaning of the text. Using the background of his own experience, the person doing the interpreting analyses the significant circumstances that have led to the linguistic statement. The important point is thus not merely that a normative statement is correctly applied to the “life situation” which is to be judged. On the contrary, according to *hermeneutics* both elements – life situation and linguistic statement – are dependent upon each other.<sup>8</sup> In this context, Gadamer speaks of an “ontological change”<sup>9</sup>; the “subject” that stands in the tradition of Descartes, Kant and Fichte is replaced by a theory of language, which however remains true to the subject-based philosophical tradition insofar as the way we understand something is still exposed to our previous understanding.<sup>10</sup>

7 In dieser Nuance anders: Karl Larenz/Claus-Wilhelm Canaris, *Methodenlehre der Rechtswissenschaft* (1995), 204.

8 Grundlegend: Hans-Georg Gadamer, *Hermeneutik. Wahrheit und Methode – Grundzüge einer philosophischen Hermeneutik*, *Gesammelte Werke* (2010). Zur Jurisprudenz vgl. unter vielen Karl Larenz, *Methodenlehre der Rechtswissenschaft*, *Enzyklopädie der Rechts- und Staatswissenschaft*. Abteilung Rechtswissenschaft (1991), 204 f; Josef Esser, *Vorverständnis und Methodenwahl in der Rechtsfindung*. Rationalitätsgrundlagen richterlicher Entscheidungspraxis (1972), 159 ff; Winfried Hassemer, *Juristische Hermeneutik*, ARSP (1986), 195–212, 195 ff.

9 Hans-Georg Gadamer, *Hermeneutik. Wahrheit und Methode – Grundzüge einer philosophischen Hermeneutik*, *Gesammelte Werke*, 385 ff.

10 Thomas Vesting, *Rechtstheorie als Medientheorie* (Supplement I) – Überlegungen zur Notwendigkeit der Verknüpfung von Sprachtheorie und Medientheorie, *Ancilla Iuris* (anci.ch) (2010), 43–84, 51.

7 Different views on this nuance: Karl Larenz/Claus-Wilhelm Canaris, *Methodenlehre der Rechtswissenschaft* (1995), 204.

8 According to Hans-Georg Gadamer, *Hermeneutik. Wahrheit und Methode – Grundzüge einer philosophischen Hermeneutik*, *Gesammelte Werke* (2010). On legal method see Karl Larenz, *Methodenlehre der Rechtswissenschaft*, *Enzyklopädie der Rechts- und Staatswissenschaft*. Abteilung Rechtswissenschaft (1991), 204 f; Josef Esser, *Vorverständnis und Methodenwahl in der Rechtsfindung*. Rationalitätsgrundlagen richterlicher Entscheidungspraxis (1972), 159 ff; Winfried Hassemer, *Juristische Hermeneutik*, ARSP (1986), 195–212, 195 ff.

9 Hans-Georg Gadamer, *Hermeneutik. Wahrheit und Methode – Grundzüge einer philosophischen Hermeneutik*, *Gesammelte Werke* (2010), 385 ff.

10 Thomas Vesting, *Rechtstheorie als Medientheorie* (Supplement I) – Überlegungen zur Notwendigkeit der Verknüpfung von Sprachtheorie und Medientheorie, *Ancilla Iuris* (anci.ch) (2010), 43–84, 51.

Interpretieren ist somit einerseits von einer verselbständigten Sprachlogik und andererseits durch Vorverständnisse geprägt. Unbeeinflusste Aussagen lassen sich vorweg nicht gewinnen. Wenn sich Rechtswissenschaftler und Rechtswissenschaftlerinnen dennoch als *Wissenschaftler* sehen, so deshalb, weil sie die Rechtstexte auf verschiedene Deutungsmöglichkeiten hin befragen und diese Abfrage so transparent wie möglich darstellen.<sup>11</sup> Ganz im Sinne der Hermeneutik steigern die Rechtswissenschaftlerinnen und -wissenschaftler so die Objektivität durch intersubjektive Verständigung innerhalb der Wissenschaft – wenn auch die Dichte des intersubjektiven Diskurses gering und der Bias des kollektiven Vorverständnisses bestehen bleibt.<sup>12</sup>

Die Steigerung der Objektivität durch intersubjektive Verständigung muss aber nicht auf einen Diskurs, welcher sich auf subjektive Deutungen durch Rechtswissenschaftlerinnen und Rechtswissenschaftler stützt, beschränkt bleiben. Die Hinwendung der vor allem US-amerikanischen Rechtswissenschaft zur *Interdisziplinarität*<sup>13</sup> und zu einer *sozialwissenschaftlichen Empirie* vermochte die intersubjektive Verständigung markant zu erweitern; sozialwissenschaftliche Empirie bedeutet hier, dass die Wissenschaft nicht nur eine, sondern in systematischer Weise eine Vielzahl von Sprachäusserungen oder eine Vielzahl mit Sprachäusserungen verknüpfte Umstände befragt, um zum gesuchten Verständnis zu gelan-

The act of interpretation is thus characterised on the one hand by an independent logic of language and on the other hand by our prior understanding. No statements can be made that are not subject to prior influence. If those study legal science nevertheless see themselves as *scientists*, this is because they question legal texts to find out what the possible interpretations are, and represent this questioning as transparently as possible.<sup>11</sup> Very much in accordance with the principles of hermeneutics, legal scientists thus increase their objectivity by intersubjective communication within their academic discipline – even if the density of the inter-subjective discourse remains at a low level and the bias of collective prior understanding continues to apply.<sup>12</sup>

However, this increase in objectivity through inter-subjective communication does not have to remain limited to a discourse based on subject-based interpretations by legal scientists. The move towards *interdisciplinarity*<sup>13</sup> and a *social sciences based empiricism* which is particularly evident among US legal scientists has brought about a striking expansion in inter-subjective communication; here, “social sciences based empiricism” means that scientists question not one, but (in a systematic way) a large number of linguistic statements, or a large number of circumstances linked to linguistic statements, in order to arrive at the understanding they are aiming for.<sup>14</sup> The success of empirical studies coincides with increased

11 Karl Larenz/Claus-Wilhelm Canaris, Methodenlehre der Rechtswissenschaft (1995), 204 f.

12 Philipp Sarasin, Diskurstheorie und Geschichtswissenschaft, in: R. Keller (Hrsg.), Handbuch für sozialwissenschaftliche Diskursanalyse. Bd. I: Theorien und Methoden (2001), 29–52; Lorenz Engi, Politische Verwaltungssteuerung – demokratisches Erfordernis und faktische Grenzen (2008), Wissenschaft und Werturteil.

13 Vgl. grundlegend Lawrence Meir Friedman, Das Rechtssystem im Blickfeld der Sozialwissenschaften (1981); Lawrence Meir Friedman/Stewart Macaulay, Law and the behavioral sciences (1969).

11 Karl Larenz/Claus-Wilhelm Canaris, Methodenlehre der Rechtswissenschaft (1995), 204 f.

12 Philipp Sarasin, Diskurstheorie und Geschichtswissenschaft, in: R. Keller (Hrsg.), Handbuch für sozialwissenschaftliche Diskursanalyse. Bd. I: Theorien und Methoden (2001), 29–52; Lorenz Engi, Politische Verwaltungssteuerung – demokratisches Erfordernis und faktische Grenzen (2008), Wissenschaft und Werturteil.

13 Cf. Lawrence Meir Friedman, Das Rechtssystem im Blickfeld der Sozialwissenschaften (1981); Lawrence Meir Friedman/Stewart Macaulay, Law and the behavioral sciences (1969).

gen.<sup>14</sup> Der Erfolg empirischer Studien fällt dabei zusammen mit der vermehrten Intervention in selbstregulierte Gesellschaftsbereiche mittels Recht: Ein liberal geprägtes Rechtssystem, welches Freiräume definiert und nur punktuell eingreift, vermag zwar selbstregulierte Gesellschaftsbereiche zu schaffen. Die heute als nötig empfundenen konstanten Nachsteuerungen und Korrekturen innerhalb dieser gesellschaftlichen Selbstregulierungen erfolgen dagegen pragmatisch, ad hoc und bedürfen ihrerseits stets der Beobachtung und Nachsteuerung.<sup>15</sup>

Da sich, nach dem hier dargelegten Verständnis, die Empirie auf dem Boden der Hermeneutik bewegt und sich damit nicht wesensmässig, aber immerhin graduell von einer traditionellen, qualitativen (Rechts-)

intervention by the law in self-regulated areas of society: although a legal system that is characterised by its liberality, defining areas of freedom and only occasionally intervening, can create self-regulated areas of society, the constant readjustments and corrections within such social self-regulation that are currently seen as necessary are on the other hand undertaken pragmatically, on an ad hoc basis, and in turn require constant observation and readjustment.<sup>15</sup>

Since empiricism, according to the understanding as represented here, forms the basis of hermeneutics, and is thus distinguished from traditional, qualitative (legal) science in terms of degree rather than by its

14 Niklas Luhmann, *Die Wissenschaft der Gesellschaft* (1992), 369 f., mit kritischer Anmerkung zu den möglichen Folgen einer resultierenden Überkomplexität. Empirie wird gemeinhin definiert als das methodische Sammeln von Daten resp. Tatsachen, auf deren Grundlage mittels Induktion Erkenntnisse gewonnen werden. Bereits Oliver Wendell Holmes sah die Zukunft der Rechtswissenschaft in der Empirie: *Oliver W. Holmes, The Path of the Law*, Harvard Law Review (1920), 457–469, 469. Während sodann erste empirische Studien durch Rechtswissenschaftler ab den 1940er Jahren erstellt wurden, setzte sich die empirische Forschung an den amerikanischen Universitäten innerhalb der letzten 10 Jahre durch: *Tracey E. George, An Empirical Study of Empirical Legal Scholarship: The Top Law Schools*, Indiana Law Journal (2006), 141–148. Seit 2004 wird das *Journal of Empirical Legal Studies* publiziert. Vgl. aus einer Europäischen Sicht hierzu jüngst *Axel Tschentscher, Rechtsvergleichung und empirische Forschung*, *Zeitschrift für vergleichende Rechtswissenschaft (ZVglRWiss)* (2009), 362–377.

15 Diese Erkenntnis formulierte erstmals Polanyi 1944 mit seiner Theorie der Doppelbewegung: *Karl Polanyi, The great transformation – politische und ökonomische Ursprünge von Gesellschaften und Wirtschaftssystemen* (1995); siehe hierzu *Marc Amstutz, Historizismus im Wirtschaftsrecht*, in: Rainer J. Schweizer/Herbert Burkert/Urs Gasser (Hrsg.), *Festschrift für Jean Nicolas Druey zum 65. Geburtstag* (2002), 9–29.

14 Niklas Luhmann, *Die Wissenschaft der Gesellschaft* (1992), 369 f., with critical note concerning the possible consequences of resulting over-complexity. Empiricism is generally defined as the methodical collection of data/facts which can be used as a basis for findings by induction. Oliver Wendell Holmes foresaw the future of legal science as based on empiricism: *Oliver W. Holmes, The Path of the Law*, Harvard Law Review (1920), 457–469, 469. While the first empirical studies by legal scientists date from the 1940s, empirical research has become well established at American universities within the last 10 years: *Tracey E. George, An Empirical Study of Empirical Legal Scholarship: The Top Law Schools*, Indiana Law Journal (2006), 141–148. The *Journal of Empirical Legal Studies* has been published since 2004. Cf on this point from a European perspective, recently: *Axel Tschentscher, Rechtsvergleichung und empirische Forschung*, *Zeitschrift für vergleichende Rechtswissenschaft (ZVglRWiss)* (2009), 362–377.

15 This realisation was first formulated by Polanyi in 1944 with his theory of the double movement: *Karl Polanyi, The great transformation – politische und ökonomische Ursprünge von Gesellschaften und Wirtschaftssystemen* (1995); according to *Marc Amstutz, Historizismus im Wirtschaftsrecht*, in: Rainer J. Schweizer/Herbert Burkert/Urs Gasser (Hrsg.), *Festschrift für Jean Nicolas Druey zum 65. Geburtstag* (2002), 9–29.

Wissenschaft unterscheidet, bleibt sie für eine qualitative Rechtswissenschaft anschlussfähig: Die qualitativen Argumente wie jene zum Wandel des Staates sind als Hypothesen vorgebracht, als Modelle<sup>16</sup> von Recht-in-Gesellschaft, oder sie lassen sich als solche formulieren. Als solche Modelle machen sie eine wissenschaftliche Debatte darüber möglich, ob die vorgebrachten Aussagen zum behaupteten staatlichen Wandel zutreffen oder nicht. Sie sind – mit empirischen Mitteln – falsifizierbar<sup>17</sup> oder lassen sich zumindest in Frage stellen. Thema ist dann, wie qualitative Argumente in Rechtssetzung, Rechtsprechung und Rechtswissenschaft zu empirisch feststellbaren Veränderungen im Recht stehen.<sup>18</sup> Die (falsifizierende) Empirie erweitert den wissenschaftlichen Diskurs auf der Grundlage von qualitativ gewonnenen Hypothesen und bietet dergestalt eine gegenüber der qualitativen Wissenschaft erweiterte Objektivität durch intersubjektive Verständigung<sup>19</sup>.

Einer der möglichen interdisziplinären und empirischen Zugänge zum Recht bietet die *Korpuslinguistik*. Der Vorteil der Korpuslinguistik liegt zunächst darin, dass sie sich als Teil der Linguistik auf die Sprache und damit direkt auf die Be-Gründung des Rechts richtet. Mit der Linguistik lässt sich

nature, it remains compatible with qualitative legal science: the qualitative arguments, such as the argument concerning how the state has changed, are submitted as hypotheses, as models<sup>16</sup> of “law-in-society”, or lend themselves to being formulated in that way. As models of this kind, they make it possible for a scientific debate to be conducted as to whether the statements submitted concerning the alleged transformation of the state are true or not. They are – using empirical means – falsifiable<sup>17</sup>, or at least open to question. The issue then is how qualitative arguments in law-making, jurisdiction and legal science stand in relation to empirically identifiable changes in the law.<sup>18</sup> (Falsifying) empiricism extends the scientific discourse on the basis of qualitatively acquired hypotheses and in this way, through intersubjective communication, offers greater objectivity by comparison with qualitative science<sup>19</sup>.

One possible interdisciplinary and empirical approach to law is provided by *corpus linguistics*. The benefit of corpus linguistics lies firstly in the fact that as a branch of linguistics it is aimed at language, and therefore directly at the justification (Be-Gründung) of law. With linguistics, it is pos-

16 Zum Modellgedanke siehe Rudolf Wiethölter, Sozialwissenschaftliche Modelle im Wirtschaftsrecht, Kritische Justiz (KJ) (1985), 126–139, m.w.H.

17 Grundlegend hierzu Karl R. Popper, The growth of scientific knowledge (1963/1979). Eine sozialwissenschaftliche Empirie resp. Falsifizierung ist freilich nicht mit jener der Naturwissenschaften gleichzusetzen: vgl. Fn. 14 und 18 f.

18 Nach Adorno gehören quantitative und qualitative Forschung zusammen: Theodor W. Adorno, Soziologie und empirische Forschung (1957), in: Gesammelte Schriften, Band 8 (1998), 196–216.

19 Zur intersubjektiven Verständigung, auf deren Boden sich auch empirische Wissenschaft immer bewegt: Jürgen Habermas, Erkenntnis und Interesse, Theorie (1968), 198.

16 Concerning the model principle see Rudolf Wiethölter, Sozialwissenschaftliche Modelle im Wirtschaftsrecht, Kritische Justiz (KJ) (1985), 126–139, m.w.H.

17 According to Karl R. Popper, The growth of scientific knowledge (1963/1979). Eine sozialwissenschaftliche Empirie resp. Falsifizierung ist freilich nicht mit jener der Naturwissenschaften gleichzusetzen: cf. Fn. 14 and 18 f.

18 According to Adorno, quantitative and qualitative research go hand in hand: Theodor W. Adorno, Soziologie und empirische Forschung (1957), in: Gesammelte Schriften, Band 8 (1998), 196–216.

19 Concerning intersubjective communication which also always forms the basis for empirical science: Jürgen Habermas, Erkenntnis und Interesse, Theorie (1968), 198.

die Frage stellen, ob behauptete normative Veränderungen in der Konzeption des Staates mit semantischen Veränderungen im Rechtsdiskurs einhergehen. Mit dieser Frage können die qualitativ vorgebrachten Hypothesen wie jene zum Wandel des Staates getestet werden, und zwar indem nach einem empirisch belegbaren Wandel des Staatsbegriffs gesucht wird und ein so festgestellter Wandel im Umgang mit dem Argument Staat den qualitativ beschriebenen Änderungen gegenübergestellt wird.

Die *Linguistik* wurde in der Rechtswissenschaft und in verwandten Themen bisher bereits vereinzelt eingesetzt:

- Eine ältere Tradition der Rechtslinguistik fokussiert auf *qualitative* Analysen von Rechtstexten, um den Unterschied von Rechtssprache und juristischer Dogmatik zur alltagsweltlichen Sprache freizulegen.<sup>20</sup>

- Aus den *staatsvergleichenden Politikwissenschaften* stammen Untersuchungen von Schmitt/Obinger, welche mit einer empirischen Verfassungsvergleichung für das öffentliche Wirtschaftsrecht die Eigenarten schweizerischen Verfassungsrechts untersuchen. Sie geben punktuell Erklärungen dafür, welchen Wegen Privatisierungsbestrebungen in der Schweiz folgen.<sup>21</sup>

sible to pose the question of whether alleged normative changes in the way we conceive of the state are associated with semantic changes in legal discourse. This question can be used to test qualitatively submitted hypotheses (concerning changes in the state, for example) by searching for an empirically verifiable change in the term “state” and setting any such identified change in the way the “state” argument is dealt with against the qualitatively described changes.

Previously, *linguistics* has only been used in legal science and related topics on an occasional basis:

- An older tradition of legal linguistics focuses on *qualitative* analyses of legal texts in order to reveal how legal language and juristic theory differ from the language of the everyday world.<sup>20</sup>

- In the *political sciences*, comparisons between states have produced studies by Schmitt/Obinger, who use an empirical comparison between constitutions to investigate the particular characteristics of Swiss constitutional law in regard to public commercial law. These studies provide some clarification concerning the directions followed by privatisation efforts in Switzerland.<sup>21</sup>

---

20 Ekkehard Felder, *Juristische Textarbeit im Spiegel der Öffentlichkeit* (2003), 14 ff.

21 Carina Schmitt/Herbert Obinger, *Verfassungsschranken und die Privatisierung öffentlicher Dienstleistungen im internationalen Vergleich*, *Politische Vierteljahresschrift* (2010), 643–664; vgl. des Weiteren Herbert Obinger, *Vergleichende Policyanalyse. Eine Einführung in makroquantitative und makro-qualitative Methoden*, in: K. Schubert/N. C. Bandelow (Hrsg.), *Lehrbuch der Politikfeldanalyse* (2009), 221–244; Niklas Potrafke, *Does Government Ideology Influence Deregulation of Product Markets? Empirical Evidence from OECD Countries*, *Public Choice* (2010), 135–155.

---

20 Ekkehard Felder, *Juristische Textarbeit im Spiegel der Öffentlichkeit* (2003), 14 ff.

21 Carina Schmitt/Herbert Obinger, *Verfassungsschranken und die Privatisierung öffentlicher Dienstleistungen im internationalen Vergleich*, *Politische Vierteljahresschrift* (2010), 643–664; vgl. des Weiteren Herbert Obinger, *Vergleichende Policyanalyse. Eine Einführung in makroquantitative und makro-qualitative Methoden*, in: K. Schubert/N. C. Bandelow (Hrsg.), *Lehrbuch der Politikfeldanalyse* (2009), 221–244; Niklas Potrafke, *Does Government Ideology Influence Deregulation of Product Markets? Empirical Evidence from OECD Countries*, *Public Choice* (2010), 135–155.

– Im deutschen Raum sind zudem verschiedene *fachterminologische Arbeiten* zu finden.<sup>22</sup> Besonders interessant sind jüngste Arbeiten von Friedemann Vogel, der mit einem diskurslinguistischen Blick (semi-) automatische Analyseverfahren der Korpuslinguistik nutzt, um wiederkehrende Sprachmuster als Spuren sedimentierter Rechtsdogmatik aufzuspüren und zu interpretieren.<sup>23</sup>

– Am nächsten an die in der Folge vorgestellte Methode kommen *amerikanische Versuche mit Korpuslinguistik*. Evans/McIntosh/Lin/Cates stellen verschiedene Möglichkeiten vor, wie mit Hilfe von Textklassifikation eine detaillierte Inhaltsanalyse von politischen und juristischen Dokumenten in grosser Zahl möglich ist.<sup>24</sup> Mouritsen untersucht, ob die Auslegung der Gerichte nach der grammatikalischen Methode tatsächlich dem Wortgebrauch entspricht, wie er in maschineller Weise aus grossen Datenmen- gen herausgelesen werden kann.<sup>25</sup>

– Various *studies on specialist terminology* have also been carried out in Germany.<sup>22</sup> Of particular interest are some recent studies by Friedemann Vogel, who adopts a discourse linguistics perspective to use (semi-) automatic analysis methods of corpus linguistics to detect and interpret recurrent language patterns as traces of sedimented legal doctrine.<sup>23</sup>

– Coming closest to the method presented in the following are *American studies using corpus linguistics*. Evans/McIntosh/Lin/Cates present various possibilities for carrying out a detailed content analysis of large numbers of political and juristic documents using text classification.<sup>24</sup> Mouritsen examines whether the interpretation of the courts according to the grammatical method really corresponds to word usage as derivable by automatic analysis of large quantities of data.<sup>25</sup>

22 Leonhard Voltmer, *Werkzeuge für Rechtsdatenbanken: Über computerlinguistische Verfahren zur Untersuchung, Speicherung und Kommunikation rechtlichen Wissens* (Diss.) (2005); Leonhard Voltmer, *Computerlinguistik für die Terminografie im Recht* (2006).

23 Friedemann Vogel, *Das Recht im Text. Rechtssprachlicher Usus in korpuslinguistischer Perspektive*, in: Ekkehard Felder/Marcus Müller/Friedemann Vogel (Hrsg.), *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen von Texten und Gesprächen* (2012), 314–353.

24 Michael Evans/Wayne McIntosh/Jimmy Lin/Cynthia Cates, *Recounting the courts? Applying automated content analysis to enhance empirical legal research*, *Journal of Empirical Legal Studies* (2007), 1007–1039.

25 Stephen C. Mouritsen, *Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning*, *Columbia Science and Technology Law Review* (2011), 156–205.

22 Leonhard Voltmer, *Werkzeuge für Rechtsdatenbanken: Über computerlinguistische Verfahren zur Untersuchung, Speicherung und Kommunikation rechtlichen Wissens* (Diss.) (2005); Leonhard Voltmer, *Computerlinguistik für die Terminografie im Recht* (2006).

23 Friedemann Vogel, *Das Recht im Text. Rechtssprachlicher Usus in korpuslinguistischer Perspektive*, in: Ekkehard Felder/Marcus Müller/Friedemann Vogel (Hrsg.), *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen von Texten und Gesprächen* (2012), 314–353.

24 Michael Evans/Wayne McIntosh/Jimmy Lin/Cynthia Cates, *Recounting the courts? Applying automated content analysis to enhance empirical legal research*, *Journal of Empirical Legal Studies* (2007), 1007–1039.

25 Stephen C. Mouritsen, *Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning*, *Columbia Science and Technology Law Review* (2011), 156–205.



### III. Konkrete Anlage einer korpuslinguistischen Untersuchung des Wandels des Staates

#### A. Methodik

Für Fragestellungen wie jener nach dem Argument des „Wandel des Staates“ können qualitative und quantitative Methoden verknüpft und die *Technik der (semi-)automatischen Korpuslinguistik* wie folgt nutzbar gemacht werden:

- *Qualitativ gemachte Aussagen* zum Wandel des Staates in ausgewählten Rechtsgebieten werden als zu testende Hypothesen ausgewählt.

- Mit der *Methode der Korpuslinguistik* wird *empirisch* untersucht, ob sich der Begriff des Staates im Laufe der Zeit gemäss Hypothese verändert hat. Der Wandel des Staates drückt sich vermutungsweise in der rechtfertigenden Beschreibung von Aufgaben und Handlungsmitteln des Staates aus. Untersuchungsgegenstand ist somit der Ausdruck des Staatsbegriffs respektive seiner Ausprägungen in juristischen Texten. Die Resultate dieser quantitativen Untersuchung werden den Hypothesen entgegengesetzt und diese gegebenenfalls falsifiziert oder in Frage gestellt.<sup>26</sup>

- Aus den empirischen Resultaten und dem Prozess der Falsifizierung werden abgewandelte oder *neue Hypothesen* gewonnen, welche wiederum empirisch untersucht werden können.

Neben der zeitlichen Gliederung können mit Hilfe von *thematischen Unterteilungen* zusätzliche Erkenntnisse gewonnen werden. Für die Frage nach dem Wandel des Staates bietet es sich an, unterschiedliche Rechtsgebiete zu befragen, um allfälligen Differenzen auf die Spur zu kommen. Zudem liegt es – mit Blick auf die verschiedenen institutionellen Aspekte des Rechts – nahe, Texte aus den Bereichen der Rechtssetzung, der Recht-

### III. Specific structure of a corpus linguistics based study of the transformation of the state

#### A. Method

For issues such as the question concerning the argument of the “transformation of the state”, qualitative and quantitative methods can be combined and the *technique of (semi-)automatic corpus linguistics* can be used as follows:

- *Qualitative statements* concerning the transformation of the state in selected areas of the law are chosen as hypotheses to be tested.

- Using the *corpus linguistics method*, an *empirical* investigation is carried out into whether the concept of the state has changed over the course of time in accordance with what the hypothesis states. The transformation of the state is assumed to be expressed in the justifying description of the state’s tasks and means of action. The subject of investigation is therefore the way in which the concept of the state or its manifestations is expressed in juristic texts. The results of this quantitative investigation are compared with the hypotheses, which will be falsified or called into question.<sup>26</sup>

- From the empirical results and the falsification process, modified or *new hypotheses* will be derived, which can in turn be subjected to empirical investigation.

In addition to classification according to periods of time, further knowledge can be gained by means of *thematic subdivisions*. As far as the question of the transformation of the state is concerned, it is useful to approach different areas of the law in order to identify possible differences. It is also clear – in light of the various institutional aspects of the law – that texts from the spheres of legislation, jurisdiction and legal

---

26 Zur Falsifizierung vgl. oben Fn. 17.

---

26 Concerning falsification, see footnote 17 above.

sprechung und der Rechtswissenschaft zu berücksichtigen und miteinander zu vergleichen. Untersuchungsgrundlage sind somit linguistisch aufbereitete Textsammlungen, welche das Gesamtkorpus darstellen. Mit der vorliegenden Untersuchung sollte primär getestet werden, ob und in welcher Weise sich die Korpuslinguistik für rechtswissenschaftliche Analysen eignet. Die Untersuchungsgrundlage, d.h. die korpuslinguistisch aufzubereitenden Texte mussten somit zwangsläufig begrenzt werden. Um dennoch eine repräsentative Auswahl an Aussagen zu erhalten, wurden drei möglichst unterschiedliche Bereiche des öffentlichen Rechts gewählt, in welchen *alle*<sup>27</sup> Texte erfasst werden sollten: Sicherheitsrecht, öffentliches Wirtschaftsrecht und Sozialrecht<sup>28</sup>. Für die vorliegende Untersuchung

science should be taken into account and compared with each other. The basis for investigation is therefore linguistically prepared collections of text which form the entire corpus. The present investigation is primarily intended to test whether and in what way corpus linguistics is suitable for legal science analyses. The basis for the investigation (i.e. the corpus of texts to be prepared) had therefore necessarily to be limited. In order nevertheless to ensure a representative selection of statements, three areas of public law were selected which are as different as possible, and in which ought to cover *all*<sup>27</sup> texts: security law, public commercial law and social law<sup>28</sup>. For the present

---

27 Dies erwies sich angesichts der relativen Überschaubarkeit des schweizerischen Rechtsdiskurses als aufwändig und anspruchsvoll, aber doch realistisch: Gesammelt wurden jene Artikel, welche im Titel einen Bezug zu den gewählten Rechtsbereichen aufweisen. Nachträgliche Stichproben haben gezeigt, dass zwischen 80 % und 95 % der relevanten Artikel erfasst wurden. Angesichts dessen, dass damit eine repräsentative Auswahl an Texten im Korpus verarbeitet wurde und die Erfassungsquote weit über jener von anderen vergleichbaren Forschungsprojekten (mit ca. 20–30 % *sample rate*) liegt, sind die Ergebnisse der Abfragen zur Sprachverwendung zuverlässig.

28 Damit setzt sich die vorliegende Untersuchung dem Vorwurf aus, nicht auch noch andere Rechtsbereiche der Prüfung unterzogen zu haben. Wichtig für eine erste, pionierhafte Forschungsanlage ist indes erstens eine ausreichende *fachliche* Breite, um einseitige Ergebnisse zu vermeiden. Die ausgewählten Bereiche sind entsprechend breit gestreut. Zudem war es bei der Auswahl der Rechtsbereiche wichtig, dass diese in der Geschichte des schweizerischen Bundesstaats weit zurückreichen, was zur Breite der *historischen* Datengrundlage beiträgt. Eine zeitliche Beschränkung der Datenmenge auf die Zeit nach dem Zweiten Weltkrieg wurde zuerst in Betracht gezogen, als willkürlich und historisch – zumindest für die Schweiz – kaum zu rechtfertigen aber verworfen.

---

27 This proved to be a time-consuming and challenging (but also realistic) task, in light of the relatively manageable extent of Swiss legal discourse: articles were collected which contained a reference to the selected legal areas in their titles. Subsequent random samples showed that between 80 % and 95 % of relevant articles were included. In light of the fact that a representative selection of texts were therefore processed in the corpus and the inclusion ratio was well above that of other comparable research projects (approx. 20–30 % *sample rate*), the results of our searches into language usage are nevertheless reliable.

28 The present investigation thus lays itself open to the criticism that other areas of the law were not included. However, for an initial, pioneering research structure it is important in the first place to have a sufficient breadth of *subject matter*, in order to avoid one-sided results. The areas selected constitute a correspondingly broad spread. In addition it was important, in the selection of legal areas, that these should extend some way back into the history of the Swiss federal state, contributing to the breadth of the *historical* data base. A time-based restriction of the data quantity to the post-war period was initially considered, but was rejected as being arbitrary and historically (at least as far as Switzerland is concerned) hardly justifiable.

wurden somit alle Texte einbezogen, welche folgende Bedingungen erfüllen:

- publizierte Entscheide des Bundesgerichts, Botschaften des Bundesrats zur Gesetzgebung des Bundes und rechtswissenschaftliche Beiträge aus den wichtigsten juristischen Zeitschriften der Schweiz<sup>29</sup>, welche sich regelmässig mit den gewählten Fachbereichen befassen;
- publiziert im Zeitraum zwischen 1874 und 2012;
- Publikationen in den Bereichen Sozialrecht, öffentliches Wirtschaftsrecht und Sicherheitsrecht.

Das von uns *erstellte Korpus „Staatsverständnis“* umfasst 6498 Texte mit insgesamt 35'977'382 Wörtern und verteilt sich wie folgt auf die drei Bereiche: Sozialrecht 3025 Texte, 13'867'462 Wörter; öffentliches Wirtschaftsrecht 1872 Texte, 12'498'750 Wörter; Sicherheitsrecht 1601 Texte, 9'611'170 Wörter.<sup>30</sup> Zusätzlich wurde auch ein Korpus französischer Texte erstellt, um gegebenenfalls einen Sprachenvergleich zu ermöglichen (1716 Texte, 8'503'350 Wörter). Die Texte wurden mittels OCR-Software maschinell erkannt. Dies ist kein fehlerfreier Prozess. Die Fehler fallen jedoch, sofern sie nicht systematisch sind, bei der vorhandenen Textmenge nicht ins Gewicht.

Im Zentrum der quantitativen Analyse steht die *Annotation der Texte* mit linguistischen Informationen: Für jedes Wort wird die Grundform (Lemma) und die Wortart (Part-of-Speech-Tag) maschinell annotiert.<sup>31</sup> Das Korpus wird fallweise in Teilkorpora

investigation, therefore, all texts that fulfilled the following conditions were included:

- published decisions of the Swiss Federal Tribunal, Swiss Federal Council dispatches concerning the legislation of the Swiss Federal Government, and legal science contributions from all legal periodicals in Switzerland<sup>29</sup> that deal with the selected specialist areas on a regular basis;
- published in the period 1874 to 2012;
- publications in the spheres of social law, public commercial law and security law.

The “understanding of the state” *corpus created* by us comprises 6498 texts (in the German language) with a total word count of 35,977,382, and is distributed over the three areas as follows: social law 3025 texts, 13,867,462 words; public commercial law 1872 texts, 12,498,750 words; security law 1601 texts, 9,611,170 words.<sup>30</sup> In addition a corpus of French texts was also created, so that if appropriate a language comparison could be carried out (1716 texts, 8,503,350 words). The texts were automatically recognised using OCR software. This is not an error-free process. However, insofar as they are not systematic the errors are not significant in the context of the volume of text.

The quantitative analysis is centred on the *annotation of the texts* with linguistic information: for each word, the basic form (lemma) and the word type (part-of-speech tag) are automatically annotated.<sup>31</sup> The corpus is divided on a case-by-case basis into

---

29 Siehe auch oben bei Fn. 27.

30 Die Texte wurden manuell den drei Bereichen zugeordnet. Ein Text kann auch in mehreren Bereichen vorkommen.

---

29 See also above, footnote 27.

30 The texts were manually allocated to the three areas. A text can also appear in more than one area.

aufgeteilt, um Unterschiede zu messen. Dies ermöglicht es, in Teilkorpora signifikant häufig auftretende sprachliche Muster zu erkennen.<sup>32</sup>

Als *Korpusmanagement und Analysesystem* wird sodann die Corpus Workbench (CWB)<sup>33</sup> verwendet. Sie erlaubt die Verarbeitung von sehr grossen Datenmengen und kann zudem über Programmierschnittstellen mit weiteren spezifischen Analysewerkzeugen ergänzt werden. In Verbindung mit dem CWB Web-Interface CQPWeb<sup>34</sup> steht zudem ein Browser-Zugang zum Korpus

subcorpora, in order to measure differences. This makes it possible to identify language patterns that occur significantly often in subcorpora.<sup>32</sup>

The Corpus Workbench (CWB)<sup>33</sup> is used as the *corpus management and analysis system*. This permits the processing of very large quantities of data, with the possibility of adding further specific analysis tools via programming interfaces. In combination with the CWB Web-Interface CQPWeb<sup>34</sup> there is also browser access to the corpus, making complex searches possible using the

31 Diese automatische Analyse leistet der TreeTagger: <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>, letzter Zugriff: 29. Februar 2016. Hierzu *Helmut Schmid*, Improvements in Part-of-Speech Tagging with an Application to German, in: Proceedings of the ACL SIGDAT-Workshopauf (1995), 47–50. Für Deutsche Texte wird das Stuttgart-Tübingen-Tagsets (STTS, Schiller et al. 1995) verwendet. Die Schritte der Datenbeschaffung und der Annotation mit den oben genannten linguistischen Kategorien finden in einem *UIMA-Framework* (Unstructured Information Management Architecture) statt. UIMA hat sich in der Computerlinguistik zu einem anerkannten Standard entwickelt und erlaubt die Einbindung beliebiger Module zur Verarbeitung und Annotation von Texten: <https://uima.apache.org>, letzter Zugriff: 29. Februar 2016.

32 Zu unterscheiden sind die Angaben als absolute Frequenzen und die Frequenz pro Mio. Wörter. Die Frequenz pro Mio. ist eine Angabe, um Frequenzen in verschiedenen Korpora, die unterschiedlich groß sind, vergleichen zu können. Aber auch da sollte man die absoluten Frequenzen im Blick behalten. Zudem muss berücksichtigt werden, ob sich die Treffer auf mehrere Texte verteilen oder aber alle im gleichen Text vorkommen. Wenn ein Wort häufig ist, aber nur in einem Text vorkommt, muss dieser Fund deswegen natürlich nicht gleich unbedeutend sein, aber er muss richtig gedeutet werden: Offensichtlich bringt dieser Text ein Wort ein, das sich anderswo nicht findet – was für die Deutung interessant sein kann.

33 <http://cwb.sourceforge.net/>, letzter Zugriff: 22. Dezember 2015.

34 <http://cwb.sourceforge.net/cqpweb.php>, letzter Zugriff: 22. Dezember 2015.

31 This automatic analysis is performed by the Tree-Tagger: <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>, last access: 29. February 2016. According to *Helmut Schmid*, Improvements in Part-of-Speech Tagging with an Application to German, in: Proceedings of the ACL SIGDAT-Workshopauf (1995), 47–50. For German texts, the Stuttgart-Tübingen-Tagsets (STTS, Schiller et al. 1995) is used. The stages of data procurement and annotation with the above mentioned linguistic categories take place in an *UIMA-Framework* (Unstructured Information Management Architecture). UIMA has become a recognised standard in computer linguistics and permits the inclusion of any number of modules for the processing and annotation of texts: <https://uima.apache.org>, last access: 29. February 2016.

32 We distinguish absolute frequencies and frequencies per million words. The frequency per million is a statement by which frequencies in different corpora of different sizes can be compared. However, the absolute frequencies also need to be kept in view. In addition it has to be borne in mind whether the hits are distributed over a number of texts or all occur in the same text. If a word appears often but only in one text, this does not of course mean that this finding is unimportant, but it has to be correctly interpreted: obviously this text often uses a word that is not found elsewhere – which can certainly be of interest as far as interpretation is concerned.

33 <http://cwb.sourceforge.net/>, last access: 22. December 2015.

34 <http://cwb.sourceforge.net/cqpweb.php>, last access: 22. December 2015.

zur Verfügung, der mittels der CQP-Abfrage-sprache komplexe Recherchen im annotierten Korpus ermöglicht.<sup>35</sup> Für die Recherche können beliebige Teilkorpora gebildet werden, um Phänomene isoliert in bestimmten Datenquellen zu untersuchen.

CQP-query language in the annotated corpus.<sup>35</sup> Any number of subcorpora can be formed for the search, in order to investigate phenomena in isolation in defined data sources.

**Menu**

- Corpus queries
  - Standard query
  - Restricted query
  - Word lookup
  - Frequency lists
  - Keywords
- User controls
  - Query history
  - Saved queries
  - Categorised queries
  - Upload a query
  - Analyse corpus
  - Create/edit subcorpora
- Corpus info
  - View corpus metadata
  - Corpus documentation
  - STTS
  - TiGer Morph
- Admin tools
  - Admin control panel
  - Corpus settings
  - Manage access
  - Manage metadata
  - Manage text categories
  - Manage annotations
  - Manage visualisations
  - Cached queries
  - Cached databases

**DE-Korpus Wandel Staatsverständnis Final: powered by CQPweb**

**Restricted Query**

Query mode:

Number of hits per page:

**Select the text-type restrictions for your query:**

text_decade	text_period	text_subtopics
<input type="checkbox"/> 1870	<input type="checkbox"/> p1 (bis 1918)	<input type="checkbox"/> [Akteneinsicht](Finanzmarktrecht)
<input type="checkbox"/> 1880	<input type="checkbox"/> p2 (1919-1945)	<input type="checkbox"/> [Alkoholgesetzgebung](Berufszulassung)
<input type="checkbox"/> 1890	<input type="checkbox"/> p3 (1946-1973)	<input type="checkbox"/> [Anlagefonds](Finanzmarktaufsicht)
<input type="checkbox"/> 1900	<input type="checkbox"/> p4 (1974-1989)	<input type="checkbox"/> [Anstaltsweisung]
<input type="checkbox"/> 1910	<input type="checkbox"/> p5 (1990-heute)	<input type="checkbox"/> [Anstaltsweisung](Diskriminierungsschutz)
<input type="checkbox"/> 1920		<input type="checkbox"/> [Anstaltspolizei](Zulassung zum Universitätsstudium)
<input type="checkbox"/> 1930		<input type="checkbox"/> [Anwälte]
<input type="checkbox"/> 1940		<input type="checkbox"/> [Anwälte](Gleichberechtigung)
<input type="checkbox"/> 1950		<input type="checkbox"/> [Anwälte](Werbeverbot)
<input type="checkbox"/> 1960		<input type="checkbox"/> [Anwälte](Wirtschaftsfreiheit)
<input type="checkbox"/> 1970		<input type="checkbox"/> [Apotheken]
<input type="checkbox"/> 1980		<input type="checkbox"/> [Apotheken](Gesundheitsberufe)
<input type="checkbox"/> 1990		<input type="checkbox"/> [Arbeitsgesetz]
<input type="checkbox"/> 2000		<input type="checkbox"/> [Arbeitsgesetz](Gesundheitschutz)
<input type="checkbox"/> 2010		<input type="checkbox"/> [Arbeitsgesetz](Gewerbeaufsicht)
		<input type="checkbox"/> [Arbeitsrecht]
		<input type="checkbox"/> [Armenrecht]
		<input type="checkbox"/> [Arztberuf]
		<input type="checkbox"/> [Arztberuf](Gesundheitsberufe)
		<input type="checkbox"/> [Arztberuf](Sozialversicherung)
		<input type="checkbox"/> [Asylrecht]

Um den Wandel der Semantik und die Verwendung des Konzeptes „Staat“ zu untersuchen, muss nun die *Menge der möglichen sprachlichen Realisierungen* des abstrakten Konzeptes „Staat“ bestimmt werden: Es geht also darum, das Wortfeld und die Formulierungsmuster, in denen sich das Konzept widerspiegelt, zu bestimmen.

In Ergänzung dazu können auch korpus-gestützte lexikographische Verfahren eingesetzt werden, um Kandidaten für sprachliche Realisierungen durch Korpusanalysen aufzufinden,<sup>36</sup> oder es können mit induktiven Methoden der Datenanalyse statistisch auffällige Lexeme (Wörter, welche sich die Grundbedeutung teilen) eruiert werden.<sup>37</sup>

To investigate how the meaning and use of the concept of “state” have changed, it is now necessary to define the *quantity of the possible linguistic realisations* of the abstract concept of “Staat [state]”, i.e. to determine the word field and the formulation models in which the concept is reflected.

In addition, corpus-supported lexicographical methods can also be used in order to reveal candidates for linguistic realisations through corpus analyses,<sup>36</sup> or statistically unusual lexemes (words which have a basic meaning in common) can be determined with inductive methods of data analysis.<sup>37</sup>

35 Stefan Evert/The OCWB Development Team, The IMS Open Corpus Workbench (CWB) CQP Query Language Tutorial, <http://cwb.sourceforge.net/documentation.php>, letzter Zugriff: 29. Februar 2016.

35 Stefan Evert/The OCWB Development Team, The IMS Open Corpus Workbench (CWB) CQP Query Language Tutorial, <http://cwb.sourceforge.net/documentation.php>, last access: 29. February 2016.

Steht das zu analysierende Wortfeld fest, werden statistische Standard-Methoden wie z.B. die *Kollokationsanalyse* eingesetzt, um für die definierten Teilkorpora und Zeitabschnitte den jeweils typischen Sprachgebrauch und damit die Semantik des Konzeptes „Staat“ zu analysieren. Da die Korpora Wortarten-annotiert sind, ist auch die Analyse bestimmter syntaktischer Konstruktionen möglich, z.B. von Adjektiv-Nomen-Komplexen wie „Sozialstaat“ resp. „sozialer Staat“.

Diese traditionellen Methoden der Korpusanalyse können durch *datengeleitete Musteranalysen* ergänzt werden. Damit tritt neben die Methode der Falsifizierung von Hypothesen eine zusätzliche induktive Methode. Es werden lexikalische (Wortform, Lemma, Wortart) und grammatikalische Informationen (Nebensatztyp, Tempus, Phrasenstruktur etc.) datengeleitet zu komplexen n-Grammen (Wortgruppen) zusammengefasst. Durch statistische Vergleiche der Frequenzen dieser komplexen n-Gramme in den zeitlich definierten Teilkorpora können Konstruktionen gefunden werden, die für ein bestimmtes Teilkorpus typisch sind. Damit wird der semantischen Prämisse Rechnung getragen, dass die Bedeutung eines Wortes von dessen Umgebung abhängt, und zugleich kann die stark

Once the word field to be analysed is established, standard corpus linguistic methods such as *collocation analysis* are used in order to analyse the typical language use for the defined subcorpora and time periods, and thus the semantics of the concept of “state”. Since the corpora are annotated with part of speech classes, the analysis of particular syntactical structures is also possible, e.g. adjective-noun combinations such as “Sozialstaat” or “sozialer Staat” [social state].

These traditional methods of corpus analysis can be complemented by *data led pattern analysis*. An additional inductive method is thus used alongside the method of the falsification of hypotheses. Lexical information (word form, lemma, part of speech class) is combined with grammatical information (clause type, tense, phrase structure etc.) on a data-led basis to form complex n-grams (multiword units). By means of statistical comparisons of the frequencies of these complex n-grams in the temporally defined subcorpora, constructions can be found that are typical for a particular subcorpus. This takes into account the semantic premise that the meaning of a word depends on its environment, and at the same time the strongly hypothesis-led perspective (whereby previously defined concepts and

- 
- 36 Vgl. Cyril Belica/Kathrin Steyer, Korpusanalytische Zugänge zu sprachlichem Usus, AUC (Acta Universitatis Carolinae), Germanistica Pragensia (2006); Stefan Engelberg/Lothar Lemnitzer, Lexikographie und Wörterbuchbenutzung (2001); Stefaniy Ptashnyk/Erla Hallsteinsdóttir/Noah Bubenhofer, Korpora, Web und Datenbanken. Computergestützte Methoden in der modernen Phraseologie und Lexikographie/Corpora, Web and Databases. Computer-Based Methods in Modern Phraseology and Lexicography (2010).
- 37 Cyril Belica, Analysis of Temporal Changes in Corpora, International Journal of Corpus Linguistics (1996), 61–73; Noah Bubenhofer, Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse, Sprache und Wissen (2009).

- 
- 36 Cf. Cyril Belica/Kathrin Steyer, Korpusanalytische Zugänge zu sprachlichem Usus, AUC (Acta Universitatis Carolinae), Germanistica Pragensia (2006); Stefan Engelberg/Lothar Lemnitzer, Lexikographie und Wörterbuchbenutzung (2001); Stefaniy Ptashnyk/Erla Hallsteinsdóttir/Noah Bubenhofer, Korpora, Web und Datenbanken. Computergestützte Methoden in der modernen Phraseologie und Lexikographie/Corpora, Web and Databases. Computer-Based Methods in Modern Phraseology and Lexicography (2010).
- 37 Cyril Belica, Analysis of Temporal Changes in Corpora, International Journal of Corpus Linguistics (1996), 61–73; Noah Bubenhofer, Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse, Sprache und Wissen (2009).

hypothesegeleitete Perspektive, bei der von vorher definierten Begriffen und Formulierungsmustern ausgegangen wird, durch eine induktive Analysemethode ergänzt werden. Dieses Verfahren wurde bereits im Bereich von Stilanalysen und korpuspragmatischen Fragestellungen erfolgreich erprobt<sup>38</sup> und ist auch in semantischer Perspektive vielversprechend.<sup>39</sup>

### **B. Die Analysemöglichkeiten im Einzelnen**

Die folgenden Analysen zeigen eine Auswahl von korpuslinguistischen Methoden, die eingesetzt werden können, um den Wandel des Staatsverständnisses besser zu verstehen. Als Beispiel soll die einleitend erwähnte Hypothese dienen, dass – vor allem nach den Terroranschlägen von New York im Jahr 2001 – ein Staatsverständnis (wieder) zunehme, nach welchem der Staat die Verantwortung für seine Gesellschaft und insbesondere deren Sicherheit wahrnehme.<sup>40</sup> Als Datengrundlage wird das deutschsprachige Korpus verwendet und, sofern nichts anderes vermerkt wird, das ganze Korpus (also alle drei Bereiche: Sozialrecht, Wirtschaftsrecht und Sicherheitsrecht) miteinbezogen.<sup>41</sup>

formulation patterns are assumed) can be complemented by an inductive analysis method. This method has already been successfully tested in the field of style analyses and corpus-pragmatic studies<sup>38</sup> and is also very promising as far as the semantic perspective is concerned.<sup>39</sup>

### **B. Analysis possibilities in detail**

The following analyses show a selection of corpus linguistics methods that can be used in order to gain a better understanding of how our concept of “state” has changed. We take as an example the hypothesis we mentioned in the introduction, i.e. that – particularly after the terrorist attacks in New York in 2001 – there has (once again) been an increase in our understanding of the “state” as bearing responsibility for society and in particular for the security of society.<sup>40</sup> The German language corpus is used as a data basis, and (unless otherwise stated) the entire corpus (i.e. all three areas: social law, commercial law and security law) is included.<sup>41</sup>

38 Noah Bubenhofer/Joachim Scharloth, Kontext korpuslinguistisch: Die induktive Berechnung von Sprachgebrauchsmustern in großen Textkorpora, in: P. Klotz/P. R. Portmann-Tselikas/G. Weidacher (Hrsg.), Kontexte und Texte. Studien zu soziokulturellen Konstellationen literalen Handelns (2010); Joachim Scharloth/Noah Bubenhofer, Datengeleitete Korpuspragmatik: Korpusvergleich als Methode der Stilanalyse, in: E. Felder/M. Müller/F. Vogel (Hrsg.), Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen von Texten und Gesprächen (2011), 195–230; *ibid.*

39 Noah Bubenhofer, Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse, Sprache und Wissen (2009); *ibid.*

40 Vgl. oben Fn. 2.

41 Die hier verfolgte These wird nicht auf das Sicherheitsrecht beschränkt, sondern für alle drei Teilbereiche – Sozialrecht, Sicherheitsrecht und öffentliches Wirtschaftsrecht – untersucht. Dies im Sinne der Hypothese, um möglichen Phänomenen auf breiter Spur nachzugehen.

38 Noah Bubenhofer/Joachim Scharloth, Kontext korpuslinguistisch: Die induktive Berechnung von Sprachgebrauchsmustern in großen Textkorpora, in: P. Klotz/P. R. Portmann-Tselikas/G. Weidacher (Hrsg.), Kontexte und Texte. Studien zu soziokulturellen Konstellationen literalen Handelns (2010); Joachim Scharloth/Noah Bubenhofer, Datengeleitete Korpuspragmatik: Korpusvergleich als Methode der Stilanalyse, in: E. Felder/M. Müller/F. Vogel (Hrsg.), Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen von Texten und Gesprächen (2011), 195–230; *ibid.*

39 Noah Bubenhofer, Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse, Sprache und Wissen (2009); *ibid.*

40 Cf footnote 2 above.

41 The thesis being pursued here is not restricted to security law, but is examined in regard to all three sub-areas together – social law, security law and public commercial law – as envisaged in the hypothesis, the aim being to follow up possible phenomena on a broad basis.

1. *Das Lexem STAAT und seine zeitliche Distribution*

Als wichtigen Indikator dafür, dass über den „Staat“ gesprochen wird, kann offensichtlich das Lemma<sup>42</sup> STAAT dienen, wobei sowohl Komposita („Staatsgewalt“) als auch Derivationen („staatlich“, „rechtsstaatlich“) in die Analyse mit einbezogen werden. Die folgende Tabelle zeigt die 50 häufigsten unterschiedlichen Lexeme<sup>43</sup> von STAAT, und zwar erstens die absolute Frequenz, d.h. das tatsächliche Vorkommen im Korpus, und zweitens die relative Verwendung, d.h. das Vorkommen pro Million Wörter<sup>44</sup>:

Lexem	Frequenz		Lexem	Frequenz	
	absolut	relativ		absolut	relativ
Staat	16993	472.3	staatsvertraglich	305	8.5
staatlich	10624	295.3	Staatenlose	287	8
staatsrechtlich	3360	93.4	Staatschutz	286	7.9
Mitgliedstaat	2525	70.2	Nachbarstaat	268	7.4
Staatsangehörige	1045	29	Herkunftsstaat	257	7.1
rechtsstaatlich	977	27.2	Staatsgewalt	242	6.7
Staatsvertrag	966	26.9	Staatshaftung	221	6.1
Vertragstaat	949	26.4	Staatskasse	213	5.9
Staatsangehörigkeit	872	24.2	Staatsverwaltung	171	4.8
Rechtsstaat	786	21.8	Staatsverfassung	164	4.6
Staats-	684	19	bundesstaatlich	158	4.4
Staatsaufgabe	601	16.7	Staatssekretariat	157	4.4
innerstaatlich	542	15.1	Staatspersonals	153	4.3
Staatsgarantie	514	14.3	Staatsgebiet	151	4.2
Staatsrecht	499	13.9	Rechtsstaatlichkeit	151	4.2
Staatsrat	493	13.7	Staatsorgan	134	3.7
Bundesstaatsrecht	463	12.9	Staatsbank	132	3.7
Heimatstaat	458	12.7	Staatsrechnung	132	3.7
Drittstaat	452	12.6	nichtstaatlich	131	3.6
Staatsanwaltschaft	426	11.8	Staatsdienst	129	3.6
staatspolitisch	413	11.5	Gliedstaat	123	3.4
zwischenstaatlich	385	10.7	Staatswesen	120	3.3
Bundesstaat	338	9.4	EU-Staat	115	3.2
Verstaatlichung	326	9.1	Staatsbürger	115	3.2
Sozialstaat	324	9	einzelstaatlich	114	3.2

1. *The lexeme STAAT [STATE] and its temporal distribution*

Obviously, the lemma<sup>42</sup> STAAT [STATE] can serve as an important indicator that the subject matter here is the “state”, with compound forms (“Staatsgewalt”) [state authority] and derivations (“staatlich” [governmental], “rechtsstaatlich” [constitutional]) also being included in the analysis. The following table shows the 50 most frequent different lexemes<sup>43</sup> of STAAT, with the absolute frequency (i.e. the actual occurrence in the corpus) being stated first, followed by the relative use, i.e. the occurrence per million words<sup>44</sup>:

42 Lemma bezeichnet die Grundform eines Wortes für eine Gruppe von Lexemen wie z.B. Staatswesen.  
43 Das Lexem (Grundform) abstrahiert von der konkreten Form (Morphologie) und der syntaktischen Funktion: also z.B. „Staat“ statt „[des] Staates“.  
44 D.h. wie häufig das Lexem pro Million Wörter auftreten würde.

42 Lemma designates the basic form of a word for a group of lexemes, for example “Staatswesen” [body politic].  
43 The lexeme (basic form) abstracted from the specific form (morphology) and the syntactical function, i.e. “Staat” [state] instead of “[des] States” [of the state].  
44 I.e. how frequently the lexeme would appear per million words.



Die Frequenzen machen deutlich, dass nur wenige Lexeme sehr häufig verwendet werden, darunter „Staat“, „staatlich“, „staatsrechtlich“ und „Mitgliedstaat“. Die meisten Lexeme treten unter zehn Mal pro Mio. Wörter auf.<sup>45</sup>

Auch wenn sich aus den Frequenzen, welche sich über den ganzen Zeitraum von 1874 bis heute erstrecken, wesensmässig keine Aussagen über die Evolution der Staatlichkeit gewinnen lassen, ergeben sich bereits *Hinweise auf dominante Themen*, welche regelmässiger als andere auf das Konzept „Staat“ rekurren:

– Zunächst fallen jene Themen als besonders frequent auf, bei welchen der Staat (wie im Völkerrecht und Bundesstaatsrecht) Subjekt<sup>46</sup> oder (wie im Ausländerrecht) direkter Bezugspunkt<sup>47</sup> ist und das Lemma STAAT deshalb besonders häufig im Text erscheint.

The frequencies make it clear that only a few lexemes are used very often, including “Staat”, “staatlich”, “staatsrechtlich” and “Mitgliedstaat”. Most lexemes occur less than ten times per million words.<sup>45</sup>

Although it is essentially not possible, on the basis of the frequencies (which extend over the entire period from 1874 to the present day), to make any statement concerning the evolution of statehood, there are already *indications of dominant themes* which recur more regularly than others in relation to the concept of “state”:

– Firstly, themes that occur particularly frequently are those in which (as in international law and federal constitutional law) the state is the subject<sup>46</sup> or (as in aliens law) is the direct point of reference<sup>47</sup>, and the lemma STAAT therefore appears particularly frequently in the text.

45 Die vielen eher niedrig frequenten Lexeme sind keine Überraschung: Allgemein folgen Wortfrequenzen in einem beliebigen Textkorpus (egal welcher Sprache) dem sog. Zipf’schen Gesetz, nach dem die Frequenz eines Wortes umgekehrt proportional zu seinem Rang in der sortierten Wortliste ist. Das bedeutet, dass es nur sehr wenige Wörter gibt, die sehr häufig und sehr viele, die nur selten vorkommen (vgl. *Rainer Perkuhn/Holger Keibel/Marc Kupietz, Korpuslinguistik* (2012), 84).

46 Mitgliedstaat, Staatsangehörige, Staatsvertrag, innerstaatlich, Vertragsstaat, Bundesstaat, zwischenstaatlich, Bundesstaatsrecht, Nachbarstaat, einzelstaatlich, bundesstaatlich, Staatsgebiet, Gliedstaat.

47 Staatsangehörige, Staatsangehörigkeit, Heimatstaat, Drittstaat, Staatenlose, Nachbarstaat, Herkunftsstaat, Staatsbürger, EU-Staat.

45 The many rather infrequently used lexemes come as no surprise: generally, word frequencies in any text corpus (irrespective of the language) follow the so called Zipf’s Law, which states that the frequency of a word is inversely proportional to its rank in the sorted word list. This means that there is only a very small number of words that occur very often and a very large number that occur very rarely (cf. *Rainer Perkuhn/Holger Keibel/Marc Kupietz, Korpuslinguistik* (2012), 84).

46 Mitgliedstaat [Member State], Staatsangehörige [citizens], Staatsvertrag [treaty], innerstaatlich [internal], Vertragsstaat [contracting state], Bundesstaat [federal state], zwischenstaatlich [interstate], Bundesstaatsrecht [federal state law], Nachbarstaat [neighbouring state], einzelstaatlich [national], bundesstaatlich [federal], Staatsgebiet [national territory], Gliedstaat [member state].

47 Staatsangehörige [nationals], Staatsangehörigkeit [nationality], Heimatstaat [home state], Drittstaat [third country], Staatenlose [stateless persons], Nachbarstaat [neighbouring state], Herkunftsstaat [state of origin], Staatsbürger [citizen], EU-Staat [EU state].

– Als besonders wichtige Themen, die mit dem Staat in engem Zusammenhang stehen, erscheinen zudem die Staatsaufgaben, das Finanzrecht (mit vermutlich primär historischen Bezügen: Staatsgarantie, Staatskasse, Staatsbank, Staatsrechnung), die Staatsgewalt, die Staatshaftung, die Staatsverwaltung (samt Staatssekretariat, Staatspersonal und Staatsdienst) sowie die Staatsverfassung.

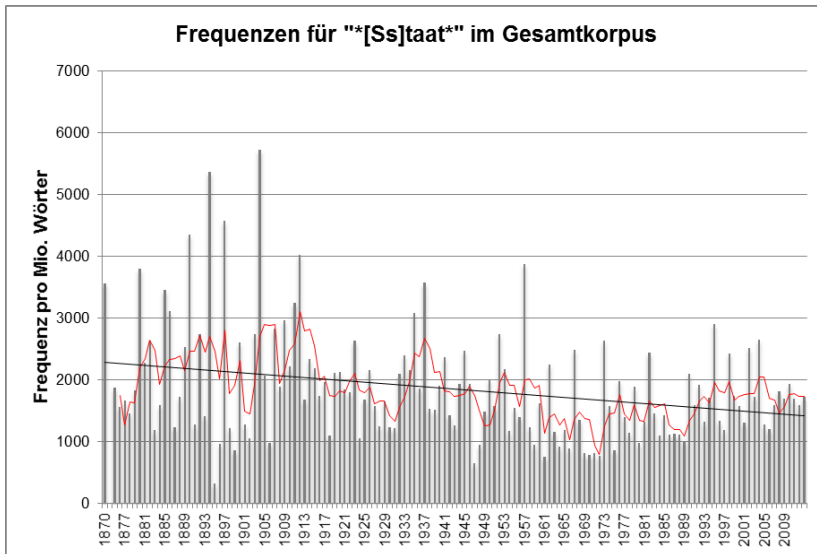
– Interessant ist zudem die sehr häufige Anrufung des Rechtsstaats (resp. der Rechtsstaatlichkeit) und des Sozialstaats. Mit der häufigen Verwendung lässt sich vermuten, dass diesen Begriffen eine besondere recht-fertigende Bedeutung zukommt. Zur Klärung dieser These wäre in einem nächsten qualitativen Schritt mit den relevanten Texten nach den konkreten Gründen für die besondere Frequenz zu fragen.

Um die Verwendung von STAATs-Lexemen in zeitlicher Abhängigkeit zu untersuchen, dient eine einfache Zeitreihe:

– Other particularly important themes closely connected with the state are the functions of the state [Staatsaufgaben], financial law (with presumably primarily historic references: Staatsgarantie [state guarantee], Staatskasse [public funds], Staatsbank [state bank], Staatsrechnung [state budget], state authority [Staatsgewalt], state liability [Staatshaftung], public administration [Staatsverwaltung] (together with state secretariat [Staatssekretariat], government personnel [Staatspersonal] and civil service [Staatsdienst]) as well as national constitution [Staatsverfassung].

– Also of interest are the very frequent invoking of the constitutional state [Rechtsstaat] (or the concept of constitutionality [Rechtsstaatlichkeit]) and the social state [Sozialstaat]. The frequent use of these terms suggests that they have a particularly justifying (recht-fertigende) significance. To clarify this thesis, the next qualitative step would be to examine the relevant texts to determine the specific reasons for this special frequency.

A simple time series serves to examine the use of STAAT lexemes in temporal dependency:



In einer derartigen Zeitreihe zum Lexem „STAAT“ ist zu erkennen, dass die Frequenz, mit welcher dieses pro Million Wörter vorkommt, von Jahr zu Jahr z.T. stark variieren kann. Ausschläge gegen oben sind in der Regel auf *Grossthemen* mit Staatsbezug zurückzuführen, welche vor allem als längere Botschaften im Bundesblatt und ausführliche Zeitschriftenartikel die Frequenz temporär nach oben treiben können. Dass z.B. das Jahr 1904 die stärkste Frequenz verzeichnet, ist massgeblich auf die Antrittsrede Max Hubers an der Universität Zürich zum Thema „Die Entwicklungsstufen des Staatsbegriffs“ zurückzuführen. Die empirische Methode der vorliegenden Untersuchung führt somit zuvorderst auf die eigene Fragestellung zurück<sup>48</sup>, oder positiv ausgedrückt: weist den Beobachter auf einen wichtigen, einschlägigen Text hin.

Die einzelnen starken Ausschläge können mit Hilfe statistischer Verfahren über die Zeit geglättet werden, sodass trotzdem eine interpretierbare Kurve über die Zeitachse entsteht. Oben werden zwei mögliche *Glättungen* in die Grafik eingeführt:

– Die Trendlinie (Gerade) folgt der linearen Regression und macht somit kenntlich, ob und in welchem Ausmass die Frequenzen über den gesamten Zeitraum einer stetig steigenden oder fallenden Geraden folgen. Im gezeigten Fall verläuft die Gerade in die neuere Zeit leicht fallend. Damit wird deutlich, dass die Frequenz, mit welcher das Lemma „STAAT“ in den untersuchten Texten zwischen 1874 und der heutigen Zeit verwendet wird, leicht abnimmt. Das Ende des Staates zu verkünden, wäre aber doch zu früh, denn bei einem fortschreitenden beständigen Abwärtstrend bliebe uns der Staat noch ca. zwei weitere Jahrhunderte erhalten.

In a time series such as this, for the lexeme “STAAT”, it is clear that the frequency with which this lexeme occurs per million words can sometimes vary greatly from year to year. Upward swings are usually attributable to *major themes* with a state reference, which – particularly in the form of lengthy Federal Council dispatches in the Federal Gazette, or detailed academic journals articles – can temporarily drive the frequency in an upwards direction. For example, the fact that the highest frequency occurs in 1904 is largely due to the inaugural speech given by Max Huber at Zurich University on the subject of “Stages in the development of the concept of the state”. Thus the empirical method used in the present study leads us straight back to the question we are considering<sup>48</sup>, or – to express it positively – points the observer in the direction of an important, relevant text.

The strong individual upward swings can be smoothed out over time with the aid of statistical methods, so that a curve arises over the time axis that is capable of interpretation. Two possible *smoothing processes* are introduced into the diagram above:

– The trend line (straight line) follows the linear regression and thus shows whether and to what extent the frequencies follow a constantly rising or falling straight line over the entire period. In the case shown, the straight line tends to fall slightly as it approaches the more recent period, clearly indicating that the frequency with which the lemma “STAAT” is used in the texts between 1874 and the present time decreases slightly. However, it would be premature to start talking of the demise of the state, for if the downwards trend were to continue at a steady rate the state would still be around for approximately another two hundred years.

48 Max Huber, Die Entwicklungsstufen des Staatsbegriffs, Zeitschrift für Schweizerisches Recht (ZSR) (1904), 1.

48 Max Huber, Die Entwicklungsstufen des Staatsbegriffs, Zeitschrift für Schweizerisches Recht (ZSR) (1904), 1.

– Die rote Kurve in obiger Grafik zeigt die Möglichkeit, die grossen Ausschläge zu mildern, indem jeweils für ein Fenster von vier Jahren Durchschnittswerte berechnet werden. Dadurch werden singuläre Ausschläge gedämpft. Hiermit wird z.B. erkennbar, dass in den Jahren zwischen 1958 und 1989 der Staat in den untersuchten Rechtsbereichen weit unterdurchschnittlich angerufen wurde, bevor das Lemma „STAAT“ sodann ab 1990 wieder merklich mehr Verwendung findet. Eine Zunahme für die Jahre nach den Terroranschlägen in den USA im Jahr 2001 lässt sich dagegen – entgegen der geäusserten Behauptung eines wieder-erstarkten Staates<sup>49</sup> – nicht erkennen.

In einem weiteren Untersuchungsgang können zusätzliche Differenzierungen vorgenommen werden, um der gewonnenen These nachzuspüren. Die nachfolgende Grafik zeigt eine Unterscheidung der Zeitreihe zum Lemma STAAT *nach der Art der Texte*: Bundesgerichtsentscheide („BGE“), Botschaften und Zeitschriftenartikel („Zeitschriften“). Dabei zeigt sich, dass die zunehmende Verwendung des Wortes „Staat“ ab ca. 1990 weitgehend dessen Verwendung in den Botschaften geschuldet ist, also mit Gesetzesprojekten zusammenhängt, während die Verwendung in den Bundesgerichtsentscheiden stabil bleibt und in der Lehre gar abnimmt.

– A red curve in the above diagram shows the possibility of evening out the swings by calculating average values over a window of four years, which reduces individual swings (called “moving average”). For example, it becomes clear that in the years between 1958 and 1989 the state is invoked with well below average frequency in the legal areas under investigation, while from 1990 onwards the lemma “STAAT” is once again used noticeably more frequently. By contrast, there is no discernible increase for the years following the terrorist attacks in the USA in 2001 – contrary to the stated allegation that the state has gained a new strength<sup>49</sup>.

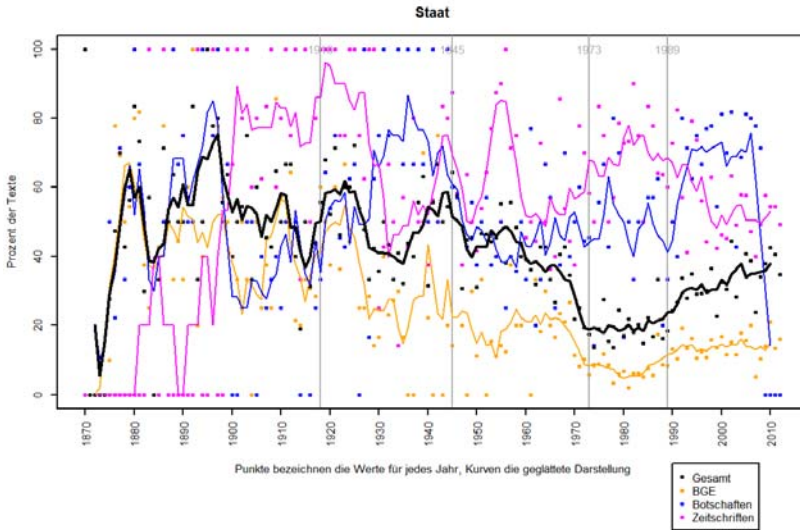
In a further investigation procedure, additional differentiations can be undertaken in order to trace the thesis obtained. The following diagram shows a differentiation of the time series for the lemma STAAT *based on the type of text*: decisions of the Swiss Federal Tribunal (Bundesgerichtsentscheide or “BGE”), Federal Council dispatches, and articles by legal scholars (“academic journals”). This shows that the increasing use of the word “Staat” from about 1990 onwards is largely due to its use in Federal Council dispatches, i.e. it is connected with legislation projects, while its use in decisions of the Swiss Federal Tribunal remains stable, and its use in legal doctrine actually declines.

---

49 Oben Fn. 2.

---

49 Cf footnote 2.



## 2. Schlagwortanalysen

Um weitere Einsichten zur Frage zu gewinnen, wie sich die Sprache im Recht über die Zeit verändert, kann sodann mit einer *Schlagwortanalyse* untersucht werden, ob die markante Zunahme des Lemmas STAAT ab ca. 1989 mit der Zu- oder Abnahme anderer Lexeme einhergeht. Mit der Schlagwortanalyse werden die Lexeme berechnet, die in einem Teilkorpus (z.B. Texte von 1989 bis heute) *statistisch signifikant häufiger als im Referenzkorpus* (im vorliegenden Fall jeweils alle Texte im Gesamtkorpus) auftreten. Es wird somit eine Übersicht über die herrschenden Themen gewonnen.

In den folgenden beiden Abbildungen werden jene Lexeme angezeigt, die für die jeweiligen zeitlichen Perioden im Vergleich zum Gesamtkorpus typisch sind. Für alle Lexeme gilt, dass der Frequenzunterschied zum jeweiligen Referenzkorpus hoch signifikant ist (LLR,  $df=1$ ,  $p < 0.001$ ; dargestellt werden jeweils maximal die 200 signifikantesten Lexeme mit einer Minimalfrequenz von 3). Die Größe des Lexems drückt die statistische Signifikanz aus: Je größer, desto signifikanter ist der Frequenzunterschied.

## 2. Key word analysis

In order to gain further insight into the question of how language as used in the law changes over time, a *key word analysis* can be undertaken to ascertain whether the striking increase in the lemma STAAT from about 1989 onwards is associated with any increase or decrease in other lexemes. With key word analysis, the lexemes that occur in a subcorpus (e.g. texts from 1989 to the present day) with a *statistically significant greater frequency than in the reference corpus* (in the present case, all texts in the entire corpus) are calculated. This provides an overview of the predominant themes.

The following two illustrations shown the lexemes that are typical for the respective time periods compared with the entire corpus. For all lexemes, it can be said that the frequency difference as compared with the respective reference corpus is highly significant (LLR,  $df=1$ ,  $p < 0.001$ ; a maximum of the 200 most significant lexemes with a minimum frequency of 3 are shown in each case). The size of the lexeme expresses the statistical significance: the bigger the lexeme, the more significant the frequency difference.

Die beiden Grafiken zeigen (wiederum als Beispiel und zur Verfolgung der genannten Hypothese<sup>50</sup>) die typischen Lexeme des Korpus der gesammelten Botschaften von 1975–1989 und jene von 1990 bis 2012. Auffallend ist dabei erstens, dass die Formalien (Artikel, Abs. etc.) zunehmend den Diskurs dominieren. Zweitens fällt auf, dass bis 1989 noch klar dominierende Themen erkennbar sind, wie insbesondere Wettbewerb (Kartellkommission, Preisüberwachung, Konsument etc.) und etwas weniger signifikant Umwelt und Energie (resp. Atomgesetz, radioaktiv, Wald, Schutz, Gewässer etc.) sowie Finanzhilfe und Ausländer. Ab 1989 dominiert dagegen das Thema Europa, wobei darüber hinaus zahlreiche generische Ausdrücke erscheinen, welche nicht einem spezifischen rechtlichen Sachthema zugeordnet werden können; die Aufmerksamkeit des Diskurses verflacht.

The two graphics show (again, as an example and in order to pursue the stated hypothesis<sup>50</sup>) the typical lexemes of the corpus of the collected Federal Council dispatches from 1975–1989 and from 1990 to 2012. The first striking feature is that the formalities (Artikel [Article], Abs. [paragraph] etc.) increasingly dominate the discourse. Secondly, it is noticeable that up to 1989 there are still clearly identifiable dominant themes, such as competition (Kartellkommission [Cartel Commission], Preisüberwachung [price monitoring], Konsument [consumer] etc.) and somewhat less significantly the environment and energy (Atomgesetz [Swiss Federal Atomic Energy Act], radioaktive [radioactive], Wald [forest], Schutz [protection], Gewässer [waters] etc.), and also financial assistance and aliens. From 1989 onwards, by contrast, the theme of Europe dominates, with numerous generic expressions also appearing which cannot be allocated to a specific legal subject topic; the focus of attention in the discourse decreases.

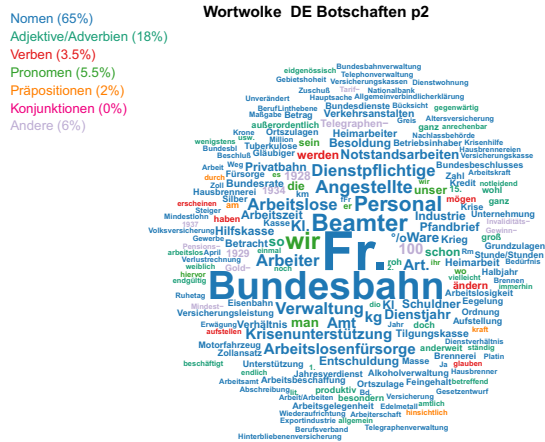


50 Vgl. oben Fn. 2.

50 Cf footnote 2 above.



A comparative glance back at the typical lexemes of the corpus of the collected Federal Council dispatches from 1919 to 1945 confirms the trends that have been identified. There are fewer signs of legal formalism in the dispatches up to 1945. In addition various highly relevant themes are very clearly shown, such as officialdom (Beamter [official], Personal [staff], Besoldung [salaries] etc.) and the Swiss Federal Railways (Verkehrsanstalten [transport agencies], Betriebsinhaber [owners of companies], Privatbahn [private railway] etc.). War is also strikingly prominent (Krisenunterstützung [crisis support], Dienstpflichtige [persons liable for compulsory service], Notstandsarbeiten [public relief works] etc.) and also the worker (Arbeitslose [unemployed persons], Arbeitszeit [working hours], Arbeitslosenfürsorge [welfare for the unemployed] etc.).



### 3. Kollokationen

Die bisher vorgestellten Analysemethoden stützen sich jeweils auf signifikant auftretende Lemma und Komposita, wobei die ganzen Texte und alle darin verwendeten Wörter als Referenzrahmen dienen. Eine weitere Möglichkeit, einem sich wandelnden Verständnis des Staates und vor allem den konkreten Themen, welche mit dem Lemma STAAT in Zusammenhang stehen, auf die Spur zu kommen, besteht darin, die jeweils typischen benachbarten Wörter (Kollokatoren) zum Lemma STAAT zu eruieren. Damit können die unterschiedlichen signifikanten Kontexte bestimmt werden, in welchen ein Lemma wie STAAT erscheint. Zusätzlich differenziert nach verschiedenen Zeitabschnitten kann hiermit eruiert werden, wie sich Beschreibungskontexte über die Zeit verändern.

Die nachfolgenden Tabellen zeigen die Kollokationen zum Lemma STAAT für die beiden Zeitabschnitte 1975–1989 und 1990 bis 2012 in allen Texten (Botschaften, BGE und Lehre) auf, und zwar jeweils die fünfzig signifikantesten Komposita.

### 3. Collocations

The analysis methods presented so far are each based on significantly occurring lemmas and compound forms, the entire texts and all the words used therein serving as a reference framework. A further possibility for identifying a changing understanding of “the state”, and in particular the specific themes that are connected with the lemma STAAT, consists in determining the words that typically appear adjacent to the lemma STAAT (collocates). This makes it possible to determine the various significant contexts in which a lemma such as STAAT appears. With additional differentiation according to different time periods, this can also be used to determine how description contexts change over time.

The following tables show the collocations for the lemma STAAT for the two periods 1975–1989 and 1990 to 2012 in all texts (Federal Council dispatches, Swiss Federal Tribunal decisions, and legal doctrine), specifically the fifty most significant compounds in each case.



Die Kollokationen<sup>51</sup> für die Periode von 1975–1989:

The collocations<sup>51</sup> for the period 1975–1989:

There are 7,725 different lemmas in your collocation database for "[word=".\*[Ss]taet.\*"]". (Your query "[word=".\*[Ss]taet.\*"]", in subcorpus "<cm>corpus\_p4</cm>" returned 6,363 matches in 520 different texts)

No.	Lemma	Total no. in this subcorpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
1	Beschwerde	2741	37.625	568	175	2139
2	Wirtschaft	820	11.256	137	36	454.17
3	Zentralblatt	44	0.604	46	42	398.65
4	vereinigt	46	0.631	44	19	361.02
5	die	489030	6712.7	8207	484	355.15
6	demokratisch	161	2.21	62	32	319.94
7	Verwaltungs	44	0.604	40	40	316.4
8	Vereinbarung	485	6.657	79	22	257.78
9	und	86583	1188.5	1742	324	233.68
10	unser	1335	18.325	112	52	225.11
11	Regeste	873	11.983	92	92	222.79
12	schweizerisch	4603	63.183	196	119	182.35
13	ausländisch	1232	16.911	95	48	176.97
14	Giacometti	62	0.851	30	18	172.32
15	Eingriff	796	10.926	74	47	162.23
16	Angehörige	244	3.349	44	13	152.65
17	Herausgeber	59	0.81	27	27	151.1
18	Intervention	143	1.963	34	19	137.78
19	Schweiz	3104	42.607	138	76	136.73
20	Flüchtling	285	3.912	43	11	133.71
21	italienisch	174	2.388	35	17	129.36
22	Bürger	628	8.62	56	35	118.58
23	in	94971	1303.6	1706	302	117.46
24	europäisch	368	5.051	44	32	116.94
25	freiheitlich	86	1.181	25	20	112.44
26	Wirtschaftspolitik	270	3.706	37	15	108.05
27	als	19697	270.37	455	173	106.7
28	Verletzung	1011	13.878	65	54	101.2
29	Bundesgericht	2926	40.164	118	89	100.88
30	Kultur-	15	0.206	13	3	99.776
31	deutsch	598	8.209	49	28	96.424
32	Recht	5023	68.949	163	101	94.322
33	Rechtsstaat	179	2.457	29	11	94.295
34	modern	237	3.253	32	21	92.52
35	Amerika	22	0.302	14	10	91.457
36	Publikation	189	2.594	29	27	91.141
37	Zürich	2074	28.469	92	65	90.824
38	Gesetzgebung	644	8.84	49	23	90.129
39	Organ	533	7.316	44	25	87.155
40	,	228680	3139	3656	384	86.51
41	jugoslawisch	13	0.178	11	3	83.246
42	Aufgabe	1916	26.3	84	58	81.522
43	Interventionismus	45	0.618	16	8	79.461
44	mit	25199	345.9	522	210	79.177
45	Demokratie	102	1.4	21	13	78.635

51 „Log-Likelihood“-Werte (letzte Spalte) bei Kollokationen sind signifikant, wenn sie größer als 3.48 sind. Werte über 10.828 sind nach sozialwissenschaftlichem Massstab „höchst signifikant“. Allerdings gilt es auch immer die absoluten Frequenzen im Blick zu haben: Eine Kollokation mag zwar signifikant sein, aber im Korpus nur zehn Mal vorkommen. Vgl. *Stefan Evert*, *The Statistics of Word Cooccurrences. Word Pairs and Collocations* (2005); *John Sinclair*, *Corpus, Concordance, Collocation* (1991); *Kathrin Steyer*, *Kollokationen in deutschen Wörterbüchern und in der deutschen Wörterbuchforschung*, *Lexicographica* (2008), 185–207.

51 “Log-Likelihood” values (last column) for collocations are significant if they are greater than 3.48. Values over 10.828 are “highly significant” according to the social sciences criterion. However, it is also important always to bear the absolute frequencies in mind: a collocation may be significant, but may only appear in the corpus ten times. Cf. *Stefan Evert*, *The Statistics of Word Cooccurrences. Word Pairs and Collocations* (2005); *John Sinclair*, *Corpus, Concordance, Collocation* (1991); *Kathrin Steyer*, *Kollokationen in deutschen Wörterbüchern und in der deutschen Wörterbuchforschung*, *Lexicographica* (2008), 185–207.

Für die Zeitperiode ab 1990 sind jene Kollokationen im Fettdruck herausgehoben, die neu erscheinen:<sup>52</sup>

For the period from 1990 onwards, collocations appearing for the first time are shown in bold:<sup>52</sup>

There are 20,537 different lemmas in your collocation database for "[word=".\*[Ss]taat.\*"]". (Your query "[word=".\*[Ss]taat.\*"]", in subcorpus "<em>corpus\_p5</em>" returned 33,141 matches in 1,835 different texts)

No.	Lemma	Total no. in this subcorpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
1	Beschwerde	8835	156.48	1351	367	3612.6
2	die	2.00E+06	32319	42196	1700	3141.1
3	<b>Regulierung</b>	1953	34.59	450	71	1576.1
4	Aufgabe	7860	139.21	740	214	1320.1
5	<b>Rechtsvorschrift</b>	407	7.208	201	64	1064.8
6	Zentralblatt	123	2.178	127	104	1032.7
7	<b>Heimat-</b>	178	3.153	143	18	978.45
8	in	390274	6912.1	9561	1283	946.14
9	<b>Private</b>	2020	35.776	331	125	928.95
10	Ander	10864	192.41	739	275	925.11
11	vereinigt	284	5.03	157	77	880.62
12	Wirtschaft	1983	35.121	318	167	878.99
13	demokratisch	860	15.231	227	121	861.3
14	<b>Schutzpflicht</b>	348	6.163	154	53	771.52
15	Staatsangehörige	700	12.398	192	70	744.68
16	<b>Handeln</b>	9422	166.87	620	276	744.42
17	schweizerisch	17632	312.28	896	463	742.35
18	und	343408	6082.1	8272	1246	737.22
19	eine	334693	5927.7	8084	1180	732.26
20	europäisch	7093	125.62	517	242	702.98
21	Verwaltungs	102	1.807	87	87	617.22
22	<b>Gemeinschaft</b>	1849	32.748	247	143	596.13
23	<b>staatlich</b>	6375	112.91	448	152	583.31
24	Angehörige	1682	29.79	234	82	582.68
25	Staat	7847	138.98	488	215	544.24
26	Recht	21756	385.32	902	423	514.43
27	<b>EU</b>	2176	38.539	240	92	494.72
28	ausländisch	4542	80.443	341	153	479.65
29	<b>EU-</b>	89	1.576	69	17	462.52
30	<b>privat</b>	5889	104.3	385	161	458.29
31	zwischen	15205	269.3	670	318	431.3
32	Amerika	87	1.541	64	33	416.63
33	<b>Selbstregulierung</b>	1087	19.252	149	33	366.81
34	modern	789	13.974	130	83	366.19
35	Demokratie	257	4.552	82	38	345.93
36	<b>Steuerung</b>	400	7.084	95	20	338.76
37	<b>Mitgliedstaat</b>	2348	41.585	204	84	335.84
38	<b>Preisordnung</b>	60	1.063	48	29	327.61
39	<b>seco</b>	368	6.518	90	71	326.58
40	<b>Staatsziel</b>	85	1.505	54	10	325.21
41	Eingriff	2908	51.503	224	127	324.29
42	als	71623	1268.5	1952	639	323.79
43	Intervention	542	9.599	101	60	309.29
44	Bürger	1195	21.165	141	89	307.84
45	<b>Privatisierung</b>	645	11.424	107	40	302.83
46	durch	44142	781.8	1303	507	295.94
47	<b>ersuchenden</b>	97	1.718	53	14	295.5

52 Entfernt wurden die Verweise auf Literaturstellen, welche aufgrund des Lexems „STAAT“ als frequent erschienen.

52 The references to citations which appear as frequent because of the lexeme “STAAT” have been removed.

Hier ist nun ein Wechsel konkreter Themen nach 1989/90 gut zu erkennen. Wenig überraschend ist zunächst, dass – im Kontext und Nachgang der EWR-Abstimmung von 1992<sup>53</sup> – Die Themen Gemeinschaft, EU, Mitgliedstaat etc. als besonders signifikant erscheinen. Dieses Ergebnis lässt sich auch bereits aus der Schlagwortanalyse ableseN.<sup>54</sup> Ebenfalls neu erscheint die Diskussion über die Schutzpflicht (des Staates). Und schliesslich tauchen viel verwendete stehende Begriffe auf, welche das Lemma STAAT beinhalten: staatliche Markt- und Preisordnungen (Art. 3 Abs. 1 lit. a KG) sowie Heimat- bzw. Herkunftsstaat (AsylG). Interessant ist, dass ein neuer Themenbereich besonders dominiert: die Regulierung (mit den verwandten Begriffen wie Steuerung, Selbstregulierung, Preisordnung, Privatisierung, Staatsziel, Eingriff, Intervention etc.). In der Rangliste der meistverwendeten Lexeme belegt „Regulierung“ denn auch den dritten Rang mit einem sehr hohen *log-likelihood*-Wert (*llr*) von 3141. Im vorangehenden Zeitabschnitt betrug dieser nur 78. Ähnlich zeigt sich das Bild, wenn man das Kompositum STAAT – Regulierung<sup>55</sup> betrachtet: Ab 1990 ist es besonders frequent mit einem *llr* von 332, während dieser zuvor nur knapp 35 und vor 1975 sogar nur 20 betrug. Interessant ist zudem, dass das Kompositum in der Zeit von 1874–1918 bereits einmal eine breitere Verwendung fand (*llr* von 48). Diese Beobachtungen stimmen mit einer Abfrage im Korpus auf die zeitliche Distribution des Lemmas REGULIER und einer (auf die deutsche Sprache begrenzte) Abfrage nach den Lexemen „Regulierung“ und „regulieren“ bei Googles *Ngram Viewer*, welcher alle Texte der von Google gescannten Bücher

A change in specific themes after 1989/90 is clearly identifiable here. It is hardly surprising, firstly, that – in the context and aftermath of the 1992 European Economic Area referendum<sup>53</sup> – the themes of community, the EU, member State etc. appear particularly significant. This result is also evident from the key word analysis.<sup>54</sup> Another new phenomenon is the discussion concerning the duty (i.e. the state's duty) of protection. And finally, there are also established and much-used concepts which contain the lemma STAAT: staatliche Markt- und Preisordnungen [governmental market and price regulations (Art. 3 (1) a of the Swiss Cartel Act [KG]) and Heimatstaat [home state] or Herkunftsstaat [state of origin] (Swiss Asylum Act). It is interesting that one new thematic area is particularly dominant: that of regulation (with the associated concepts of control, self-regulation, price regulation, privatisation, state goal, interference, intervention etc.). In the ranking for the most used lexemes, “regulation” occupies 3<sup>rd</sup> place with a very high *log-likelihood* ratio (*llr*) of 3141. In the preceding time period, the *llr* was only 78. The picture is similar when we look at the compound form STAAT – Regulierung<sup>55</sup>: from 1990 onwards it occurs particularly frequently, with an *llr* of 332, compared with just 35 before 1990 (and before 1975 only 20). It is also interesting that the compound form was already widely used in the period 1874–1918 (with an *llr* of 48). These observations accord with a query in the corpus for the temporal distribution of the lemma REGULIER and a query (restricted to the German language) regarding the lexemes “Regulierung” and “regulieren” in Google's *Ngram Viewer*, which combines all texts of books that have been

53 Der EWR-Beitritt der Schweiz scheiterte bei der Volksabstimmung am 6. Dezember 1992.

54 Oben Seite 23.

55 Z.B. „staatliche Regulierung“ oder „Staatsregulierung“.

53 The national referendum on 6 December 1992 put paid to Switzerland's entry into the EEA.

54 Cf page 23 above.

55 For example “staatliche Regulierung” or “Staatsregulierung” [both meaning governmental regulation].

vereinigt, überein. Die Benutzung von Googles *Ngram Viewer* ist indes aus verschiedenen Gründen kritisch zu betrachten, da die Datengrundlage nicht transparent ist. Trotzdem ist die grosse Datenmenge geeignet, um generelle Trends im allgemeinen Sprachgebrauch zu prüfen.<sup>56</sup> Für detailliertere Analysen sind spezialisierte Referenzkorpora zu benützen, wie das deutsche Referenzkorpus (DeReKo, IDSMannheim)<sup>57</sup> oder die DWDS-Korpora der Berlin-Brandenburgischen Akademie.<sup>58</sup> Die erwähnten Tendenzen sind denn auch im Kernkorpus der DWDS-Korpora sichtbar.<sup>59</sup>

scanned by Google. However, for various reasons the use of Google's *Ngram Viewer* is to be viewed critically, since the data basis is not transparent. In spite of this the large quantity of data is suitable for examining general trends in general usage.<sup>56</sup> For more detailed analyses, specialised reference corpora should be used, such as the German reference corpus (DeReKo, IDSMannheim)<sup>57</sup> or the DWDS corpora of the Berlin-Brandenburg Academy.<sup>58</sup> The aforementioned trends are also clearly evident in the DWDS corpora.<sup>59</sup>

56 Jean-Baptiste Michel/Yuan Kui Shen/Aviva Presser Aiden/Adrian Veres/Matthew K. Gray/The Google Books Team/Joseph Pickett/Dale Hoiberg/Dan Clancy/Peter Norvig/Jon Orwant/Steven Pinker/Martin A. Nowak/Erez Liebermann Aiden, *Quantitative Analysis of Culture Using Millions of Digitized Books*, Science (2011), 176–182.

57 Marc Kupietz/Cyril Belica/Holger Keibel/Andreas Witt, *The German Reference Corpus DeReKo: A primordial sample for linguistic research* (2010).

58 Vgl. zum Ganzen Lothar Lemmitzer/Heike Zinsmeister, *Korpuslinguistik. Eine Einführung* (2006), 107 ff.

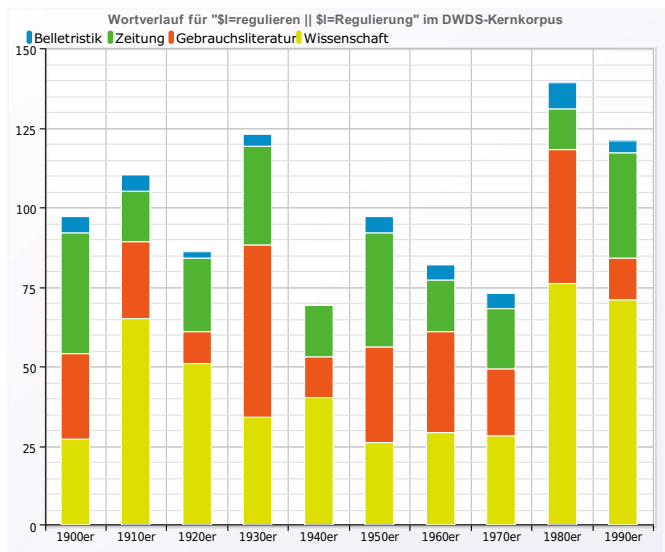
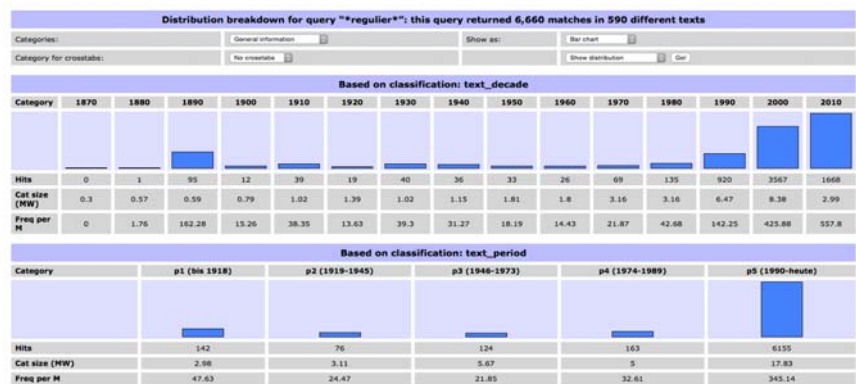
59 Das DWDS-Kernkorpus vereinigt für jede Dekade des 20. Jahrhunderts eine ähnliche Zahl von Texten aus den Domänen Belletristik, Zeitungen, Gebrauchsliteratur und Wissenschaftstexte und eignet sich deshalb besonders gut für diachrone Abfragen. Über alle Domänen hinweg sind wie in Googles *ngram*-Korpus Spitzen in den 1910er, 1930er und dann ab den 1980er Jahren sichtbar. Besonders ausgeprägt sind in den wissenschaftlichen Texten die Spitzen in den 1910er und dann ab den 1980er Jahren. Leider reicht das Korpus gegenwärtig nur bis ins Jahr 2000, sodass die weitere Entwicklung noch nicht ablesbar ist. Vgl. zum Korpus [www.dwds.de](http://www.dwds.de), Alexander Geyken, *The DWDS corpus: A reference corpus for the German language of the 20th century*, in: C. Fellbaum (Hrsg.), *Collocations and Idioms: Linguistic, lexicographic, and computational aspects* (2007), 23–42.

56 Jean-Baptiste Michel/Yuan Kui Shen/Aviva Presser Aiden/Adrian Veres/Matthew K. Gray/The Google Books Team/Joseph Pickett/Dale Hoiberg/Dan Clancy/Peter Norvig/Jon Orwant/Steven Pinker/Martin A. Nowak/Erez Liebermann Aiden, *Quantitative Analysis of Culture Using Millions of Digitized Books*, Science (2011), 176–182.

57 Marc Kupietz/Cyril Belica/Holger Keibel/Andreas Witt, *The German Reference Corpus DeReKo: A primordial sample for linguistic research* (2010).

58 Vgl. zum Ganzen Lothar Lemmitzer/Heike Zinsmeister, *Korpuslinguistik. Eine Einführung* (2006), 107 ff.

59 The DWDS core corpus combines, for each decade of the 20th century, a similar number of texts from the domains of fictional writing, newspapers, functional literature and scientific texts, and is therefore particularly suitable for diachronic queries. Over all domains, as in the Google *ngram* corpus, there are peaks in the 1910s, the 1930s and then from the 1980s onwards. The peaks in the 1910s and from the 1980s onwards are particularly pronounced in the scientific texts. Unfortunately the corpus currently only goes up to 2000, so that further developments are not shown. Cf concerning the corpus [www.dwds.de](http://www.dwds.de), Alexander Geyken, *The DWDS corpus: A reference corpus for the German language of the 20th century*, in: C. Fellbaum (Hrsg.), *Collocations and Idioms: Linguistic, lexicographic, and computational aspects* (2007), 23–42.



Insoweit kann festgestellt werden, dass der Staat im Rechtsdiskurs erstens nicht (resp. nicht nachweisbar) erst infolge der Terroranschläge im Jahr 2001 in New York häufiger angerufen wurde, sondern eine markante Zunahme in der Verwendung des Wortes „Staat“ bereits ab 1990 im Zusammenhang mit den Grossthemen Europa und Regulierung erfolgt. Insofern ist mit der empirischen Linguistik eine Hypothese gefunden, welche mittels qualitativen Mitteln weiter zu prüfen ist: Dass der Staat ab 1990 wieder mehr in die recht-fertigenden Texte Eingang fand, muss denn auch nicht zwingend als ein Erstarken des Staates interpretiert werden. In den 1990er Jahre fand unter dem Titel der Regulierung auch eine Privatisierungsdebatte statt, welche ebenfalls zu einem verstärkten Gebrauch des Lemmas „STAAT“ geführt hat (Position 45 in der oben abgedruckten Tabelle).

Kollokationsanalysen lassen sich weiter verfeinern. So ist es zum Beispiel möglich, die Analyse auf *Adjektive*, welche im Zusammenhang mit dem Lemma STAAT genannt werden, zu beschränken. Wird die Suche auf attributive Adjektive beschränkt, können sich Eigenschaften offenbaren, die dem Staat zugeschrieben oder normativ gefordert werden.<sup>60</sup> Die Resultate sind die Folgenden.

Thus it is clear that the state in legal discourse firstly did not (or did not demonstrably) only begin to be invoked more frequently in consequence of the terrorist attacks in New York in 2001; on the contrary, there has been a striking increase in the use of the word “state” from 1990 onwards in connection with the important topics of Europe and regulation. To that extent a hypothesis has been found, with the aid of empirical linguistics, which we must examine further using qualitative means: for the fact that from 1990 onwards the state once again increasingly found its way into justifying (recht-fertigende) texts does not necessarily have to be interpreted as a sign of the increasing strength of the state. In the 1990s there was also a debate about privatisation in connection with regulation, and this also led to an increased use of the lemma “STAAT” (item 45 in the above table).

Collocation analyses can be further refined. For example, it is possible to restrict the analysis to *adjectives* which are used in connection with the lemma STAAT. If the search is restricted to attributive adjectives, characteristics can be revealed which are allocated to the state or are required as standard.<sup>60</sup> The results are as follows.

---

60 Bei dieser Methode kann es natürlich sein, dass der adjektivische Kollokator nicht unbedingt im syntaktischen Bezug zum Nomen „Staat“ stehen muss, da die berechneten Kollokatoren in einem definierten „Fenster“ von fünf Wörtern links und rechts um „Staat“ erscheinen. Das Adjektiv kann sich also auch auf ein anderes Nomen, ggf. ausserhalb des Fensters, beziehen. Stichproben ergaben jedoch, dass dies selten der Fall ist und deshalb vernachlässigt werden kann.

---

60 With this method it is of course possible that the adjectival collocate does not absolutely have to stand in a syntactical relation to the noun “state”, since the collocators calculated appear in a defined “window” of five words to the left and right of “Staat”. The adjective can therefore also relate to another noun, which may be outside of the window. However, random tests have shown that this is seldom the case and can therefore be ignored.

Periode von 1975–1989:

Period from 1975–1989:

There are 10,007 different words in your collocation database for "[word=".\*staat.\*"%c]". (Your query "\*staat\*", restricted to "<em></em>: <em>p4</em>", returned 6,368 matches in 520 different texts)

No.	Word	Total no. in this subcorpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
1	Schweizerisches	331	2.728	83	66	428.01
2	Vereinigten	46	0.379	42	19	375.99
3	demokratischen	99	0.816	40	24	251.32
4	Italienische	48	0.396	20	16	127.21
5	Europäischen	239	1.97	25	21	83.312
6	westlichen	34	0.28	13	9	79.882
7	ausländischer	202	1.665	22	16	75.052
8	freiheitlichen	53	0.437	14	11	73.811
9	christlichen	17	0.14	10	7	73.056
10	westeuropäischen	25	0.206	11	10	71.5
11	türkische	7	0.058	7	6	67.179
12	modernen	121	0.997	16	10	60.769
13	militärischen	70	0.577	13	1	58.513
14	ausländische	316	2.605	21	13	51.984
15	Deutscher	40	0.33	10	7	51.481
16	demokratische	40	0.33	10	8	51.481
17	Europäischer	21	0.173	8	7	49.081
18	jugoslawische	5	0.041	5	3	47.988
19	föderativen	15	0.124	7	6	46.589
20	föderalistischen	36	0.297	9	9	46.335
21	deutsche	176	1.451	15	11	44.066
22	ausländischen	619	5.102	26	21	43.615
23	italienischen	101	0.833	12	6	43
24	jugoslawischen	6	0.05	5	2	42.586
25	geborene	114	0.94	12	12	40.135
26	liberalen	70	0.577	10	6	39.547
27	erhobene	210	1.731	15	14	39.113
28	zwischenstaatliche	40	0.33	8	2	37.274
29	abweichende	194	1.599	14	12	36.77
30	österreichische	42	0.346	8	6	36.438
31	sozialen	611	5.036	23	12	34.486
32	deutschen	342	2.819	17	13	33.338
33	Vereinigte	11	0.091	5	5	32.923
34	schweizerischen	2501	20.614	50	35	30.205
35	allgemeine	747	6.157	24	19	30.057
36	verheirateten	64	0.528	8	1	29.478
37	betreffenden	552	4.55	20	14	28.772
38	bisherige	465	3.833	18	2	27.795
39	belgischen	17	0.14	5	2	27.588
40	moderne	60	0.495	7	6	24.83

Periode von 1990–2012:

Period from 1990–2012:

There are 27,554 different words in your collocation database for "[word=".\*staat.\*"%c]". (Your query "*\*staat\**", restricted to "<em></em>: <em>p5</em>", returned 33,141 matches in 1,835 different texts)

No.	Word	Total no. in this subcorpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
1	Schweizerisches	1951	20.732	340	261	1320.1
2	Vereinigen	225	2.391	148	75	1057.8
3	demokratischen	442	4.697	123	83	602.08
4	europäischen	5010	53.239	272	152	459.79
5	modernen	407	4.325	81	65	336.97
6	ersuchenden	97	1.031	50	13	321.08
7	ausländische	1502	15.961	114	62	258.84
8	informalen	36	0.383	31	2	252.85
9	liberalen	157	1.668	42	31	201.83
10	sicheren	231	2.455	46	9	191.42
11	zuständigen	2414	25.653	121	60	188.59
12	ausländischen	2125	22.582	112	73	183.77
13	freiheitlichen	113	1.201	34	24	172.49
14	begrenzte	252	2.678	43	38	165.02
15	ersuchende	64	0.68	26	9	150.66
16	ausländischer	725	7.704	60	38	145.64
17	fremden	202	2.147	35	23	135.43
18	betreffenden	2028	21.551	93	65	131.67
19	osteuropäischen	22	0.234	17	12	131.03
20	einzelnen	5842	62.08	170	122	128.75
21	verschiedenen	3904	41.486	132	94	126.72
22	westlichen	74	0.786	24	15	125.95
23	schweizerischen	9212	97.892	227	156	125.56
24	türkische	37	0.393	19	17	121.81
25	jugoslawische	20	0.213	15	9	113.95
26	zurückgelegten	61	0.648	21	8	113.18
27	ersuchten	89	0.946	23	12	108.74
28	privater	933	9.915	55	38	100.54
29	westeuropäischen	33	0.351	16	15	100.07
30	demokratische	235	2.497	30	27	97.55
31	beherrschte	70	0.744	19	10	91.922
32	informale	13	0.138	11	2	88.865
33	voller	246	2.614	28	1	84.768
34	verfolgungssicheren	11	0.117	10	5	84.207
35	föderalistischen	86	0.914	19	19	83.29
36	verfügungsfreien	15	0.159	11	9	82.665
37	jugoslawischer	9	0.096	9	7	81.807
38	gesteuerter	12	0.128	10	3	80.121
39	eigenen	2701	28.702	88	62	79.931
40	private	2764	29.372	89	56	79.402



Wenig überraschend dominiert die Beschreibung des Staates als ein schweizerischer und ein demokratischer Staat. Ebenfalls in beiden Zeitabschnitten erscheinen der freiheitliche sowie der föderative und föderalistische Staat. Die schweizerische Ausprägung des Staates als demokratisch und freiheitlich wird somit im Rechtsdiskurs bis heute aufrechterhalten. Ebenfalls in beiden Epochen offenbart sich zudem eine Auseinandersetzung mit den „Vereinigten“ Staaten von Amerika sowie den europäischen resp. westeuropäischen Staaten.

Nimmt man wiederum die *Veränderungen von vor und nach 1989* in den Blick, so erscheint vor 1989 vor allem eine Auseinandersetzung mit einem christlichen, türkischen, militärischen und deutschen Staat. Nach 1989 fallen alle diese Adjektive aus der Rangliste, und es dominieren nun neue Beschreibungen des Staates als ein informeller, sicherer, zuständiger, begrenzter, ersuchender, osteuropäischer Staat. Dass die Sicherheit an Aktualität zunimmt, stellt dann immerhin eine eingeschränkte Analyse in Aussicht, welche die Dekade von 1990–1999 mit den nachfolgenden Jahren vergleicht: Der sichere Staat erscheint in der neueren Dekade häufiger<sup>61</sup>. Ein Blick in die konkreten Texte (resp. deren „Snippets“<sup>62</sup>) offenbart allerdings, dass dieses Resultat dem Ausländerrecht zuzuschreiben ist, welches die feste Wendung des „sicheren Drittstaates“ hervorgebracht hat:

Hardly surprisingly, the description of the state as “Swiss” or “democratic” predominates. Similarly, the “liberal” as well as the “federal” and “federalist” state appears in both periods. The Swiss characterisation of the state as “democratic” and “liberal” is thus maintained in legal discourse up to the present day. Similarly, in both periods the subject of the “United” States of America and the European (or western European) states is addressed.

If we then turn our attention to the *changes that occurred before and after 1989*, we find that before 1989 it is mainly the subject of the Christian, Turkish, military or German state that is addressed. After 1989 all of these adjectives disappear from the ranking, and new descriptions of the state as an informal, secure, responsible, limited, requesting, eastern European state predominate. A more restricted analysis comparing the decade 1990–1999 with the following years would nevertheless offer the prospect of security becoming a more topical issue: the secure state appears more frequently in the later decade<sup>61</sup>. However, a glance at the specific texts (or “snippets”<sup>62</sup> thereof) reveals that this result is attributable to aliens law, which has created the “secure third state” as an established concept:

61 Suche nach: “\*staat\* <<5>> \*sicher\*”: 1990: 363 matches in 110 different texts (in 6,723,446 words [1,116 texts]; frequency: 53.99 instances per million words). 2000: 687 matches in 214 different texts (in 8,849,319 words [1,420 texts]; frequency: 77.63 instances per million words). 2010: 191 matches in 76 different texts (in 3,139,388 words [476 texts]; frequency: 60.84 instances per million words).

62 Ein Snippet ist ein Auszug aus einem Text, der in den Suchergebnissen erscheint. Bekannt wurde der Begriff aufgrund der Google-Suche.

61 Searching for: “\*staat\* <<5>> \*sicher\*”: 1990: 363 matches in 110 different texts (in 6,723,446 words [1,116 texts]; frequency: 53.99 instances per million words). 2000: 687 matches in 214 different texts (in 8,849,319 words [1,420 texts]; frequency: 77.63 instances per million words). 2010: 191 matches in 76 different texts (in 3,139,388 words [476 texts]; frequency: 60.84 instances per million words).

62 A snippet is an extract from a text which appears in the search results. The term became known as a result of the Google search.

Your query `"staat" <3> sicheren`, restricted to `"text_decade: 2000"`, returned 42 matches in 6 different texts (in 8,375,532 words (1,361 texts); frequency: 5.01 instances per million words) (2.4 seconds)

No	Filename	Solution 1 to 42	Page 1 / 1
1	6241	Als, la ( neu ) a. Ist der Asylsuchende aus einem sicheren Drittstaat	in die Schweiz einreist, wird auf ein Asylgesuch nicht eingegangen.
2	6241	Drittstaatenangelegenheit, wie sie die Initiative vorsieht, die Bereitschaft des sicheren Drittstaats	voraus, die über sein Territorium in die Schweiz eingereisten Personen zurück
3	6241	Person zurück zu nehmen, ist die Widerspruchswie in den durchgegangenen sicheren Drittstaat	jedoch nicht gewährleistet und kann die betroffene Person nicht in den Heimat-
4	6241	Insbesondere soll auf Asylgesuche von Personen, welche über einen sicheren Drittstaat	in die Schweiz eingereist sind, künftig nicht mehr eingegangen werden.
5	6241	effektive Rückführung einer wogegenwärtigen Person in einen auf der Pflicht durchgegangenen sicheren Drittstaat	die Bereitschaft desselben zur Rückübernahme voraussetzt. Diese kann trotz der bestehenden
6	6241	verfügt, dass auf Gesuche von Asylsuchenden, die aus einem sicheren Drittstaat	in die Schweiz eingereist sind, nicht eingegangen wird, wenn der
7	6241	Landweg illegal in die Schweiz eingereist ist und damit zwangsläufig einen sicheren Drittstaat	durchgequert hat, wo sie hätte um Asyl nachsuchen können. Faktisch
8	6241	25. November 1997 erlassen, welche Richtlinien für die Bezeichnung eines sicheren Drittstaats	aufrufen. Sicheren Drittstaaten müssen danach Gewähr bieten für die Einhaltung der
9	6241	Anwendung bestimmter Konzepte und Begriffe, wie zum Beispiel für den sicheren Drittstaat	vor. Jedem Mitgliedstaat ist es frei überlassen, ob er diese
10	6241	kann die Vereinbarkeit mit der zukünftigen definitiven EU-Regelung im Bereich des sicheren Drittstaats	jedoch nicht abschliessend bestimmt werden. Die spezifischen Fragen hinsichtlich der Verteilung
11	6241	unter den europäischen Ländern waren, insbesondere bezüglich des Grundgesetzes des sicheren Drittstaats	- bereits Gegenstand verschiedener Interventionen auf unserem Kontinent. 8 Schlussfolgerungen Die
12	6254	Asylsuchende, die sich vor der Einreichung ihres Asylgesuchs in einem sicheren Drittstaat	aufgehalten haben und effektiv dorthin zurückkehren können, in diesen Staat weggewiesen
13	6254	Schutz anderweitig finden können, in einem beschleunigten Verfahren in einem sicheren Drittstaat	weggewiesen, zu welchem die asylsuchende Person eine gewisse Beziehung hat und
14	6254	Asylsuchende, die sich vor der Einreichung ihres Asylgesuchs in einem sicheren Drittstaat	aufgehalten haben und dorthin zurückkehren können, in diesen Staat weggewiesen werden
15	6254	und die Möglichkeit, dort Schutz zu finden. Unter die sicheren Drittstaaten	können auch Staaten fallen, in denen Asylsuchende, bevor sie in
16	6254	Drittstaaten besitz. Die neue Regelung macht eine klare Unterscheidung zwischen sicheren Drittstaaten	im Allgemeinen und von Bundesrat als sicher bezeichneten Drittstaaten. Bei den
17	6254	eine sehr hohe Beweislaster für die Anordnung der Weggewisung in einem sicheren Drittstaat	- muss doch ein zureichendiger Aufenthalt im Drittstaat nachgewiesen werden. Auch
18	6254	bei Nichterkenntnisentscheidungen in der Empfangsstaat - insbesondere bei Weggewisungen in einem sicheren Drittstaat	- weiterhin sichergestellt werden kann, ist ein neuer Ausschussfähigkeitsbestand vorgesehen (
19	6254	welche vorhanden ist. Die grundsätzliche Kriterien für die Bestimmung von sicheren Drittstaaten	sollen aber in der vorliegenden Berichtsfassung erläutert werden. Die Bestimmung von
20	6254	aber in der vorliegenden Berichtsfassung erläutert werden. Die Bestimmung von sicheren Drittstaaten	wird sich grundsätzlich auf europäische Staaten, insbesondere unsere Nachbarstaaten, beschränken

Dass die Sicherheit nach den Terroranschlägen von 2001 im Recht vermehrt als zentrale Staatsaufgabe thematisiert würde, lässt sich somit mit den getätigten Abfragen nicht belegen. Das bedeutet freilich nicht, dass nicht doch eine solche Wendung stattgefunden haben könnte. Wer eine solche Wendung indes trotz fehlender empirischer Nachweise behauptet, hat einen erhöhten Begründungsaufwand zu leisten.<sup>63</sup>

#### 4. Topic Models

Mit den Schlagwortanalysen und der Suche nach Kollokationen resp. entsprechenden Vergleichen über verschiedene Zeitfenster hinweg liess sich ermitteln, dass neue Themen wie Europa und Regulierung dazu geführt haben, dass das Wort „Staat“ ab den 1990er Jahren in den Texten des Korpus wieder vermehrt verwendet wurde. Eine weitere, stärker datengeleitete Möglichkeit, die Entwicklung dominierender Themen in der Zeitachse zu verfolgen, stellen die sogenannten Topic Models dar: Mit einer LDA-Klassifikation<sup>64</sup> lassen sich die Bedeutungsaspekte eines Wortes datengeleitet berechnen und anschliessend die Veränderungen der Anteile dieser Bedeutungsaspekte über

The queries submitted do not therefore provide any evidence for the idea that following the 2001 terrorist attacks security was increasingly discussed in the law as a central function of the state. Of course this does not mean that such a change of direction could not nevertheless have occurred. However, anyone asserting (in spite of the lack of empirical proof) that this change did in fact occur will have to make increased efforts to justify this view.<sup>63</sup>

#### 4. Topic Models

Using key word analyses and the search for collocations (or corresponding comparisons) over different periods of time, it has been determined that new topics such as “Europe” and “regulation” have led to the word “state” being used with increasing frequency again in the corpus texts from the 1990s onwards. An additional, more strongly data-led possibility for pursuing the development of predominant topics in the time axis is provided by the so called Topic Models: with an LDA classification<sup>64</sup> the meaning aspects of a word can be calculated on a data-led basis, and the changes in the proportions of these meaning aspects over the time axis can then be observed. All Topic

<sup>63</sup> Dass solches durchaus sinnvoll zu bewältigen ist zeigt Sibylle Hofer, Freiheit ohne Grenzen?: privatrechtstheoretische Diskussionen im 19. Jahrhundert (2001).

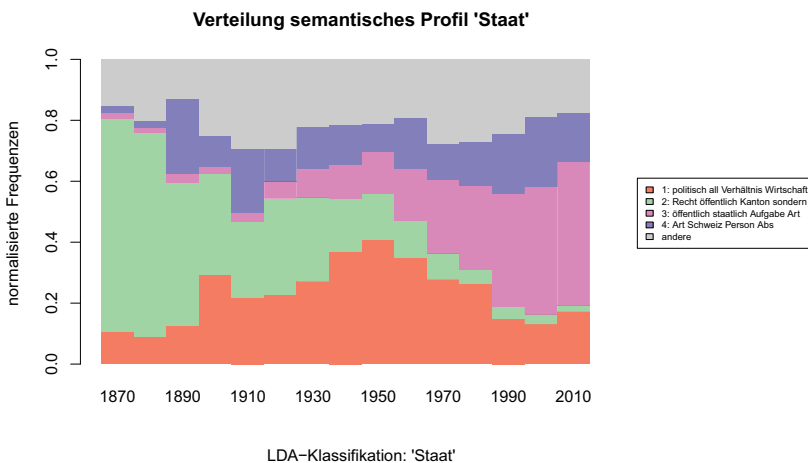
<sup>64</sup> The fact that this can certainly be meaningfully accomplished is shown by Sibylle Hofer, Freiheit ohne Grenzen?: privatrechtstheoretische Diskussionen im 19. Jahrhundert (2001).

die Zeitachse beobachten. Allen Topic Modelling-Verfahren ist gemeinsam, dass Dokumente oder Kontexte bezüglich ihres Wortmaterials maschinell in Cluster geteilt werden. Im hier gezeigten Verfahren werden Verwendungsweisen eines Wortes – 50-Wort-Textauschnitte, in denen das Wort „Staat“ erscheint – nach Ähnlichkeit geclustert.<sup>65</sup>

Im vorliegenden Korpus resultieren folgende vier Cluster von typischen Lexemen um das Wort „Staat“ (weitere Aspekte mit geringer Bedeutung werden unter „andere“ zusammengefasst):

Modelling methods have in common the fact that documents or contexts are automatically divided into clusters in relation to their word material. In the method shown here, the ways in which a word is used – 50-word text extracts in which the word “state” appears – are clustered according to similarity.<sup>65</sup>

In the present corpus the following four clusters of typical lexemes based around the word “state” appear (other aspects of lesser importance are summarised under “other”):



64 Die “Latent Dirichlet Allocation” (David M. Blei/Andrew Y. Ng/Michael I. Jordan, Latent Dirichlet Allocation, Journal of Machine Learning Research JMLR (2003), 993–1022) gehört in die Familie der Topic Modelling-Verfahren und wird besonders auch in den Digital Humanities breit angewandt. Vgl. dazu auch: *Graham Shaw/Scott Weingart/Ian Milligan, The Programming Historian: Getting Started with Topic Modelling and MALLET*, auf, <http://programminghistorian.org/lessons/topic-modeling-and-mallet> 2013, letzter Zugriff: 22. Dezember 2015.

65 Vgl. Christian Rohrdantz/Annette Hautli/Thomas Mayer/Miriam Butt/Daniel Keim/Frans Plank, Towards tracking semantic change by visual analytics, in: A. f. C. Linguistics (ed.), 49th Annual Meeting of the Association for Computational Linguistics (2011), 305–310.

64 The “Latent Dirichlet Allocation” (David M. Blei/Andrew Y. Ng/Michael I. Jordan, Latent Dirichlet Allocation, Journal of Machine Learning Research JMLR (2003), 993–1022) gehört in die Familie der Topic Modelling-Verfahren und wird besonders auch in den Digital Humanities breit angewandt. Vgl. dazu auch: *Graham Shaw/Scott Weingart/Ian Milligan, The Programming Historian: Getting Started with Topic Modelling and MALLET*, auf, <http://programminghistorian.org/lessons/topic-modeling-and-mallet> 2013, last access: 22. Dezember 2015.

65 Cf. Christian Rohrdantz/Annette Hautli/Thomas Mayer/Miriam Butt/Daniel Keim/Frans Plank, Towards tracking semantic change by visual analytics, in: A. f. C. Linguistics (ed.), 49th Annual Meeting of the Association for Computational Linguistics (2011), 305–310.

Die Cluster werden jeweils durch die typischsten Wörter (sog. Keywords) charakterisiert (in der Abbildung sind nur die ersten vier Keywords angegeben). Bei der Interpretation der Cluster dienen diese als Hilfe, um die Bedeutungsnuancen des Ausgangswortes („Staat“) zu beschreiben.

– Das Cluster 1 wird durch folgende Lexeme charakterisiert (sog. Keywords): „politisch all Verhältnis Wirtschaft damit sozial wirtschaftlich welche Schweiz Gesellschaft sondern gut ändern Frau verschieden Land einzeln heute führen“. Mit diesem Cluster zeigt sich ein Diskurs, in welchem es um das *Verhältnis von Politik und Wirtschaft zum Recht* geht, wobei als wichtige Aspekte der Sozialstaat und die Stellung der Frau erscheinen. Dieses Cluster erfährt eine erste Spitze in den 1900er Jahren und sodann seinen Höhepunkt in den 1940er und 1950er Jahren. Danach verliert dieser Diskurs an Bedeutung, scheint aber jüngst neu belebt zu werden.

– Das ermittelte Cluster 2 enthält die Keywords „Recht öffentlich Kanton sondern welche Gesetz Gemeinde Art Rechte gegenüber Beamter Konzession Interesse stehen Bank allgemein selbst solch da“. Dominierendes Thema scheint hier das *Verhältnis vom Staat zu den Kantonen, Gemeinden und Beamten* zu sein. Als wichtige Aspekte erscheinen in diesem Zusammenhang Konzessionen und Banken. Dieses Cluster dominiert zu Beginn und nimmt sodann in der Bedeutung stetig ab.

– Das Cluster 3 enthält die Keywords „öffentlich staatlich Aufgabe Art Private privat Recht Interesse Unternehmen Grundrecht Frage Bereich BV sondern Tätigkeit ob

The clusters are each characterised by the most typical words (the “key words”) (in the diagram, only the first four key words are given). In the interpretation of the clusters, these key words help to describe the meaning nuances of the initial word (“state”).

– Cluster 1 is characterised by the following lexemes (so called key words): “politisch [political] all [all] Verhältnis [relationship] Wirtschaft [economy] damit [so that] sozial [social] wirtschaftlich [economic] welche [which] Schweiz [Switzerland] Gesellschaft [society] sondern [instead] gut [good] ändern [change (verb)] Frau [woman] verschieden [different] Land [country] einzeln [individual] heute [today] führen [lead (verb)]”. With this cluster, a discourse is revealed in which the subject is the *relationship of politics and economics to the law*; the social state and the position of women appear to be important aspects. This cluster experiences an initial peak in the 1900s and then reaches a high point in the 1940s and 1950s. Thereafter, this discourse becomes less important, although it seems to have been reinvigorated recently.

– Cluster 2 contains the key words “Recht [law] öffentlich [public] Kanton [canton] sondern [instead] welche [which] Gesetz [law] Gemeinde [municipality] Art [type] Rechte [rights] gegenüber [vis-à-vis] Beamter [official] Konzession [licence] Interesse [interest] stehen [stand] Bank [bank] allgemein [generally] selbst [even] solch [such] da [there]”. The predominant theme here appears to be the *relationship between the state and the cantons, municipalities and officials*. “Licences” and “banks” appear as important aspects in this connection. This cluster is dominant at the start and then steadily becomes less important.

– Cluster 3 contains the key words “öffentlich [public] staatlich [state (adjective)] Aufgabe [task] Art [type] Private [private persons] privat [privately] Recht [law]

rechtlich grundsätzlich damit“. Dieses Cluster entspricht einem Diskurs, der die *Aufgaben des Staates im Verhältnis zu den Privaten und Unternehmen* thematisiert, wobei die Grundrechte eine wichtige Rolle spielen. Dieser Diskurs nimmt stetig und vor allem mit den 1990er Jahren an Bedeutung zu. Dies entspricht der bereits ermittelten Bedeutung des Regulierungsdiskurses, der ab den 1990er Jahren für den Bedeutungszuwachs des Lemmas „STAAT“ mitverantwortlich zeichnet.<sup>66</sup>

– Das Cluster 4 enthält die Keywords „Art Schweiz Person Abs ausländisch Recht schweizerisch welche Ausländer Angehörige Fall Bestimmung fremd Verordnung Gebiet betreffend ändern solch andere“. Dieses errechnete Cluster fokussiert auf das Ausländerrecht. Es erfährt einen markanten Höhepunkt in den 1910er Jahren und sodann in jüngerer Zeit eine zunehmende Bedeutung.

Interessant ist an diesen Clusteranalysen, dass *datengeleitet ermittelte dominante Themen zueinander in Beziehung und in eine zeitliche Dimension gesetzt* werden können. Dabei wird die bereits ermittelte These gestützt, dass der Regulierungsdiskurs in jüngster Zeit den Staat als recht-fertigendes Argument dominiert. Zudem zeigt sich, dass diese gestiegene Aufmerksamkeit für Regulierung und für das Verhältnis gegenüber Privaten und Unternehmen mit einer sinkenden Aufmerksamkeit für staatsorganisatorische Themen einhergeht. Das Thema Sicherheit insbesondere im Nachgang der

Interesse [interest] Unternehmen [company] Grundrecht [fundamental right] Frage [question] Bereich [area] BV [federal constitution] sondern [instead] Tätigkeit [activity] ob [whether] rechtlich [legally] grundsätzlich [in principle] damit [so that]“. This cluster corresponds to a discourse on the subject of the *tasks of the state in relation to private individuals and companies*; fundamental rights play an important role. This discourse becomes steadily more important, particularly from the 1990s onwards. This corresponds to the already established importance of the discourse on regulation, which from the 1990s onwards is jointly responsible for the growth in importance of the lemma “STAAT”.<sup>66</sup>

– Cluster 4 contains the key words “Art [type] Schweiz [Switzerland] Person [person] Abs [paragraph] ausländisch [foreign] Recht [law] schweizerisch [Swiss] welche [which] Ausländer [aliens] Angehörige [persons belonging] Fall [case] Bestimmung [determination] fremd [alien] Verordnung [ordinance] Gebiet [territory] betreffend [relating to] ändern [change (verb)] solch [such] andere [other]“. This cluster focuses on aliens law. It experiences a noticeable high point in the 1910s and then becomes increasingly important in recent times.

An interesting aspect of these cluster analyses is that *dominant topics that have been investigated on a data-led basis can be placed in relation to each other and in a temporal dimension*. This supports the already identified thesis that in very recent times the state is dominated by the discourse on regulation as justifying (recht-fertigendes) argument. It is also clear that the increased attention paid to regulation and to the relationship with private individuals and businesses goes hand in hand with a decrease in the attention paid to topics relating to state organisation. The topic of security, particularly in the after-

66 Oben S. 28.

66 Cf page 28.

Terroranschläge in den USA im Jahr 2001 vermag in dieser sehr weiten Untersuchungsperspektive nicht aufzuschneiden. Allenfalls mag man eine gewisse Wiederbelebung des Ausländerdiskurses als Anzeichen dafür deuten. Um diese These weiter zu prüfen, müssten indes weitere korpuslinguistische und qualitative Untersuchungen einsetzen.

##### 5. Weitere Analysemöglichkeiten

Neben den hier gezeigten Analysekategorien existieren eine Reihe weiterer Möglichkeiten der maschinellen Textanalyse, um den Veränderungen des Rechtsdiskurses nachgehen zu können. Die oben gezeigten Analysemethoden waren auf semantische Veränderungen des Staatsbegriffs fokussiert. Die Fragestellung lässt sich dabei natürlich erweitern und beispielsweise nach *Unterschieden in den verschiedenen Rechtsbereichen oder zwischen den Akteuren* fragen. Damit können die gezeigten Methoden kontrastiv eingesetzt werden, etwa indem Kollokationsprofile in verschiedenen Teilkorpora miteinander verglichen werden.

Die gezeigten Analysen gingen zudem vom Lexem „Staat“ (und Komposita) aus. Es ist offensichtlich, dass sich Diskurse zum Staatsverständnis nicht allein in diesem Lexem manifestieren. Deshalb bietet es sich an, das zu untersuchende Wortfeld zu erweitern. Dies kann hypothesengeleitet geschehen, indem das *Wortfeld systematisch um Synonyme, Hyper- und Hyponyme (Ober-/Unterbegriffe) etc. angereichert* wird, oder aber datengeleitet: So können beispielsweise *Schlagwortanalysen* genutzt werden, um das spezifische Vokabular thematisch einschlägiger Texte zu bestimmen. Aber auch eine Erweiterung auf komplexere Formulierungsmuster ist möglich, indem sog. Mehrworteinheiten, die für bestimmte Themen, Sachbereiche oder zeitliche Perioden typisch sind, berechnet werden.<sup>67</sup> Weiter gibt es Ansätze, Ontologien zu nutzen, um die im Text auftretenden Lexeme maschinell

math of the terrorist attacks in the USA in 2001, does not appear at all in this very broad examination perspective. Possibly, a certain revival of the discourse concerning aliens could be interpreted as a manifestation of security as a topic. To examine this thesis further, however, additional qualitative and corpus linguistics studies would have to be set up.

##### 5. Further analysis possibilities

In addition to the analysis categories shown here, there exist various other automated text analysis possibilities for tracing changes in legal discourse. The analysis methods demonstrated above focused on semantic changes in the concept of the state. The subject of study can of course be extended, for example to include the question of *differences between the different areas of the law or between players*. The methods demonstrated can thus be used on a contrastive basis, for example by a comparison between collocation profiles in different subcorpora.

In addition, the analyses shown were based on the lexeme “Staat” (and compound forms thereof). Obviously, discourses on the subject of our understanding of “state” will not be manifested in this lexeme alone. It therefore appears appropriate to extend the word field to be investigated. This can be done on a hypothesis-led basis, by *systematically enriching the word field to include synonyms, hypernyms and hyponyms etc.*, or on a data-led basis: for example, *key word analyses* can be used to determine the specific vocabulary of thematically relevant texts. However, the inclusion of more complex formulation models is also possible, whereby so called multiword units that are typical for certain topics, subject areas or time periods are calculated.<sup>67</sup> Other approaches involve the use of ontologies for automated semantic classification of lexemes appearing in the text, and then using these classifications as a

semantisch zu klassifizieren und auf der Basis dieser Klassifikationen Veränderungen abstrakter „Frames“, also typischer Deutungsrahmen, zu berechnen.<sup>68</sup>

Zahlreiche weitere Möglichkeiten werden gegenwärtig und wurden bereits im Rahmen korpus- und computerlinguistischer Paradigmen oder aber auch im Rahmen der Digital Humanities erprobt. Allerdings ist zu betonen, dass zu Beginn des Forschungsprozesses eine genaue Formulierung der Forschungsfragen und Hypothesen nötig ist, um anschliessend ein empirisch-quantitatives Messinstrument zu entwickeln, mit dem die Hypothesen geprüft werden können. Der quantitative Analyseschritt ist zudem im ganzen Forschungsprozess nur einer der notwendigen Schritte – letztlich müssen die Analyseergebnisse im Sinne einer empirisch unterfütterten Hermeneutik gedeutet werden.<sup>69</sup>

#### IV. Zusammenfassung

Die interdisziplinäre Forschung im Bereich der Rechtswissenschaft hat in den letzten Jahren an Breite gewonnen.<sup>70</sup> Rechtswissenschaftliche Forschung im Verbund mit empirisch gestützter Korpuslinguistik

basis for calculating changes in abstract “frames”, i.e. typical interpretation frameworks.<sup>68</sup>

Numerous further possibilities are currently being tested (and already have been tested) in the context of corpus linguistics and computational linguistics paradigms, and also in the context of the Digital Humanities. It should however be stressed that at the start of the research process the research questions and hypotheses must be precisely formulated so that an empirically quantitative measurement instrument can be developed by which the hypotheses can be tested. In addition the quantitative analysis stage is just one of the necessary stages in the entire research process – ultimately the analysis results have to be interpreted on the basis of empirical hermeneutics.<sup>69</sup>

#### IV. Summary

Interdisciplinary research in the field of legal science has become more broadly based in recent years.<sup>70</sup> However, no legal science research has been carried out in Switzerland in conjunction with empirically

67 Siehe hierzu Noah Bubenhofer/Joachim Scharloth, Korpuspragmatische Methoden für kulturanalytische Fragestellungen, in: N. Benitt/C. Koch/K. Müller/L. Schüler/S. Saage (Hrsg.), Kommunikation Korpus Kultur: Ansätze und Konzepte einer kulturwissenschaftlichen Linguistik, Giessen Contributions to the Study of Culture (2013), 47–66 und Noah Bubenhofer, Kollokationen, n-Gramme, Mehrworteinheiten, in: K. Roth/M. Wengeler/A. Ziem (Hrsg.), Handbuch Sprache in Politik und Gesellschaft, Sprachwissen (2015).

68 Vgl. hierzu Joachim Scharloth/David Eugster/Noah Bubenhofer, Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn, in: D. Busse/W. Teubert (Hrsg.), Linguistische Diskursanalyse. Neue Perspektiven (2013), 345–380.

69 Vgl. oben Fn. 8.

70 Siehe hierzu Paolo Becchi (Hrsg.), Interdisziplinäre Wege in der juristischen Grundlagenforschung (2007).

67 According to Noah Bubenhofer/Joachim Scharloth, Korpuspragmatische Methoden für kulturanalytische Fragestellungen, in: N. Benitt/C. Koch/K. Müller/L. Schüler/S. Saage (Hrsg.), Kommunikation Korpus Kultur: Ansätze und Konzepte einer kulturwissenschaftlichen Linguistik, Giessen Contributions to the Study of Culture (2013), 47–66 und Noah Bubenhofer, Kollokationen, n-Gramme, Mehrworteinheiten, in: K. Roth/M. Wengeler/A. Ziem (Hrsg.), Handbuch Sprache in Politik und Gesellschaft, Sprachwissen (2015).

68 Cf Joachim Scharloth/David Eugster/Noah Bubenhofer, Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn, in: D. Busse/W. Teubert (Hrsg.), Linguistische Diskursanalyse. Neue Perspektiven (2013), 345–380.

69 See footnote 8 above.

70 Cf also Paolo Becchi (Hrsg.), Interdisziplinäre Wege in der juristischen Grundlagenforschung (2007).

fehlt aber in der Schweiz vollständig. Erste interessante Ansätze sind immerhin in Deutschland und den USA zu verzeichnen. Die vorgestellte Forschungsmethode erprobt die *interdisziplinäre Zusammenarbeit* zwischen Rechtstheorie und Rechtsgeschichte mit der empirischen Linguistik. In dieser interdisziplinären Zusammenarbeit liegt die erste Bedeutung der vorgestellten Forschung.

Die rechtswissenschaftliche Forschung hat sich in der Schweiz, im Gegensatz zu den USA, bis heute der empirischen Forschung weitgehend verschlossen. Die zweite Bedeutung der vorgeschlagenen Forschungsmethode liegt entsprechend in der *Verbindung von empirischer und qualitativer Methodik*, und zwar nicht im Sinne von pseudo-naturwissenschaftlicher Forschung, sondern im Sinne einer kritischen sozialwissenschaftlichen Methode.

*Inhaltlich* lässt sich die vorgestellte Forschungsmethode sehr *breit einsetzen*. Ansatzweise und im Sinne eines Beispiels wurde hier mittels empirischer Korpuslinguistik geprüft, welche Bedeutung dem deskriptiven Argument „Wandel des Staates“ in normativer Hinsicht zukommt. Verwendet wurde ein Korpus, das aus folgenden Texten zum schweizerischen Recht besteht: Botschaften des Bundesrats, Bundesgerichtsent-scheide und Zeitschriftenartikel von 1874 bis 2012 im Bereich der Themen des öffentlichen Sicherheitsrechts, des Sozialrechts und des öffentlichen Wirtschaftsrechts.

Am Beispiel der These, dass der Staat seit 2001 wieder vermehrt die Verantwortung gegenüber seiner Gesellschaft übernimmt, wurde mit Hilfe der empirischen Linguistik gezeigt, dass das Argument „Staat“ im vorliegenden Korpus nicht erst seit 2001, sondern bereits seit 1991 vermehrt angerufen wird. Dass der Staat seit den Terroranschlägen von 2001 vermehrt als „ver-

based corpus linguistics, although some initial interesting approaches have emerged in Germany and the USA. The research method presented here explores *interdisciplinary cooperation* between legal theory and legal history, on the one hand, and empirical linguistics on the other hand. The primary significance of the research presented lies in this interdisciplinary cooperation.

In Switzerland (by contrast with the USA) legal science research has to date remained largely closed to empirical research. Accordingly, the second important aspect of the proposed research method is to be found in the *combination of empirical and qualitative methodology*, not in the sense of pseudo-scientific research but in the sense of a critical social-scientific method.

*In terms of content*, the research method presented can be *used very widely*. Here, as an initial approach and by way of example, we have used empirical corpus linguistics to consider what significance can be ascribed to the descriptive argument “transformation of the state” in a normative respect. The corpus used has been taken from the following texts relating to Swiss law: dispatches of the Swiss Federal Council, decisions of the Swiss Federal Tribunal and academic journals articles from 1874 to 2012 bearing on the topics of public security law, social law and public commercial law.

On the example of the thesis that since 2001 the state has once again increasingly been assuming responsibility towards society, it has been shown with the aid of empirical linguistics that the “state” argument has been increasingly invoked in the corpus not just since 2001, but from as long ago as 1991. On the basis of the investigations carried out, there is no empirical evidence to sup-



antwortlicher“ Staat erscheint oder eine Rückkehr eines „starken“ Staates zu verzeichnen wäre, liess sich mit den getätigten Untersuchungen empirisch nicht belegen.

Von besonderer Bedeutung ist schliesslich, dass mit der vorgestellten Forschungsmethode *zahlreiche neue Forschungsfragen* entstehen und sich rasch und mit einfachen Mitteln grob erschliessen lassen. Mit der vorliegend verfolgten Hypothese zeigt sich zum Beispiel, dass der Rechtsdiskurs auf das Argument „Staat“ nicht, wie gemäss zugrundeliegender Hypothese erwartet, nach 2001, sondern in den Jahren nach 1990 vermehrt Bezug nimmt, und zwar im Rahmen des Europadiskurses und im Bereich der Regulierung.

port the thesis that since the terrorist attacks of 2001 the state appears increasingly as “responsible”, or that there is any return to the idea of a “strong” state.

Finally, a particularly important aspect of the research method presented is that *many new research questions* emerge which can quickly be generally explored using simple means. For example, in the case of the hypothesis being pursued here it is clear that it is not (as expected on the basis of the underlying hypothesis) after 2001, but in the years following 1990 that we find legal discourse increasingly making reference to the “state” argument, specifically in the context of the “Europe” discourse and in the sphere of regulation.